



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

문학석사 학위논문

정신 인과 문제에 대한 실체
다원론적 해결책

2016년 2월

서울대학교 대학원

철학과 서양철학전공

전 승 태

정신 인과 문제에 대한 실체 다원론적 해결책

지도교수 강 진 호

이 논문을 문학석사 학위논문으로 제출함
2015년 10월

서울대학교 대학원
철학과 서양철학전공
전 승 태

전승태의 석사 학위논문을 인준함
2016년 2월

위 원 장 _____ (인)

부위원장 _____ (인)

위 원 _____ (인)

국문초록

정신 인과 문제에 대한 최근의 지배적인 답변은 비환원적 물리주의라는 입장으로 요약될 수 있다. 비환원적 물리주의는 정신적 유형 혹은 속성이 물리적 유형 혹은 속성으로 환원될 수 있음을 부정하면서도 물리주의를 고수하는 입장으로, 그 안에서는 개별자 동일성은 인정하는 부류와 개별자 동일성마저 부정하는 입장으로 크게 나뉠 수 있다. 개별자 동일성을 인정하는 전자는 결국 정신적 속성에 관한 유형 부수현상론과 다를 바 없지 않느냐는 비판을 받아왔으며, 이 비판으로부터 벗어나기 위해 개별자 동일성마저 부정하는 후자가 제안되었다고 볼 수 있다. 하지만 개별자 동일성을 부정하는 비환원적 물리주의는 정신적 개별자가 물리적 개별자와 구별된다고 주장하면서도 동시에 정신적 개별자가 물리적 개별자와 구별되는 인과력을 가지지는 않는다고 말해야 하는 곤란한 상황에 처하게 되는 것처럼 보인다. 속성들의 개별화가 그것들이 가진 인과력들에 따라 이루어진다는 것은 자연스럽게 받아들여질 수밖에 없으며, 따라서 비환원적 물리주의는 정신 인과 문제에 대한 만족스러운 답변을 주지 못한다.

비환원적 물리주의의 실패는 그것이 자명한 것으로 받아들이고 있는 물리 영역의 인과적 폐쇄성이라는 원리 때문이다. 인과적 폐쇄성을 받아들이는 한, 개별자 수준에서도 물리적이지 않은 다른 속성들은 모두 독립적인 인과력을 가지기가 어렵게 된다. 따라서 우리가 온전하게 비물리적인 속성들의 인과력을 인정하고자 한다면, 인과적 폐쇄성이 포기되어야 할 것이다. 비환원적 물리주의자들은 인과적 폐쇄성이 결코 포기될 수 없을 정도로 자명하다고 믿기 때문에 어쩔 수 없이 위에서 본 것과 같은 불만족스러운 해결책을 내놓는 것이다.

하지만 인과적 폐쇄성은 결코 자명하지 않으며, 우리가 정신 인과 문제에 대한 완전히 만족스러운 해결책을 제시하고자 한다면 거부되어야 한다는 것이 이 논문의 주장이다. 특히 나는 낸시 카트라이트 등이 과학 철학에서 이룩한 선구적인 작업들로부터 우리가 인과적 폐쇄성을 가정하는 것

보다 가정하지 않는 편이 현재 과학의 이론과 실천을 더욱 잘 설명할 수 있다는 점을 배울 수 있다고 생각한다. 정말로 그렇다면 우리가 인과적 폐쇄성을 무리하게 고수할 필요 역시 사라질 것이다.

인과적 폐쇄성을 거부하고 정신적 속성의 근본성을 받아들인다면, 우리에게는 정신 인과 문제에 대한 폭넓은 선택지들이 주어지게 된다. 그들 가운데서 나는 창발론의 일종으로 이해될 수 있는 아리스토텔레스식의 질료형상론적 실체 다원론이 다른 주요 대안들에 비해 여러 가지 면에서 더욱 만족스럽다는 점을 보여주고자 하였다.

주요어 : 정신 인과, 비환원적 물리주의, 인과적 폐쇄성, 다원론, 창발론, 카트라이트, 아리스토텔레스

학 번 : 2013-22786

목 차

서론	1
1. 정신적인 것의 무법칙성과 개별자 물리주의	5
1.1. 정신적인 것의 무법칙성	5
1.2. 개별자 동일성 기준	7
1.3. 일반주의 대 단칭주의	12
1.4. 유형 부수현상론	21
1.5. 결론	28
2. 비환원적 물리주의와 배제 문제	29
2.1. 배제 문제	29
2.2. 체계적인 과잉결정의 가능성 혹은 불가능성	32
2.3. 내용 외재주의와 배제 문제	37
2.4. 결론	43
3. 물리 영역의 인과적 폐쇄성 원리	45
3.1. 근본 법칙의 조건부적 성격	46
3.2. 모델을 통한 설명	53
3.3. 조건부 법칙에서 역량에 관한 서술로	62
3.4. 결론	71
4. 속성 이원론과 실체 이원론, 그리고 실체 다원론	73
4.1. 속성 이원론에서 실체 이원론으로	74
4.2. 반론과 응답	76
4.3. 순수 실체 이론 대 혼합 실체 이론	80
4.4. 아리스토텔레스주의적 실체 이론과 창발론	84
4.5. 결론	90
참고문헌	91
Abstract	97

서론

정신 인과 문제는 우리가 세계에 대해 가지고 있는 두 가지 그림들 사이에서 일어나는 것처럼 보이는 충돌을 해결하려는 문제이다. 한편으로 우리는 자유로운 행위자로서 우리가 믿고 원하는 바가 인과적 효력을 가지고 있다는 사실을 자명하게 여기며 살아간다. 법 제도를 비롯한 사회적 실천들은 모두 우리가 행위자라는 사실을 전제함으로써 이루어진다. 하지만 다른 한편으로 근대 과학이 제시하는 것처럼 보이는 그림에 따르자면, 모든 인과적 작용들이 근본적으로는 미시 물리적 수준에서만 이루어지는 것 같다. 이 그림은 사실 정신의 인과적 효력뿐만 아니라, 미시 물리적이지만 모든 거시적 대상들이 가지고 있는 것처럼 보이는 인과력을 위협한다. 인간이 아닌 거시적 대상들은 인간처럼 자유로운 행위자로서 인과력을 행사하는 것은 아니지만, 우리는 거시적 대상들이 그것들이 만들어내는 변화의 원인이 된다는 생각을 가지고 있다. 인과가 미시 물리적 수준에서만 이루어진다면, 인간을 포함한 모든 거시적 대상들은 진정한 원인으로서의 지위를 위협받게 될 것이다. 따라서 정신 인과 문제는 심신 문제를 넘어서서 일반화되는 성격을 지닌다.

그렇다면 이 충돌하는 두 그림들을 어떻게 할 것인가? 나는 기본적으로 우리가 행위자로서 가지고 있는 그림이 더욱 강력하고 확고한 것이라고 생각한다. 실천이 이론에 선행하기 때문이다. 이 그림은 너무도 근본적이라 과연 우리가 이것을 의심한다는 것이 정말로 말이 되는지조차 의심스러울 정도이다. 만약 두 그림 중에 하나가 포기되어야만 한다면, 근대 과학이 보여주는 것 같은 그림이 포기되어야 할 것이다. 비록 그것이 자연 속에서 인간이 위치가 어디인지에 관한 불안감을 증폭시킬지라도 그러하다. 그런 점에서 현재 이 문제에 관해 주류적인 입장인 여러 물리주의들은 근본적으로 불만족스럽다.

다행히도 나는 그저 행위자로서의 우리가 가지고 있는 그림의 실용적 중요성을 강조하는 것 말고도 더 나은 방식으로 문제를 해결할 수 있다고

믿고 있다.¹⁾ 내가 보기에 근대 과학이 제시하는 것 같은 그림은 우리가 과학의 이론과 실천을 잘못 해석한 결과로 얻게 된 오도적인 그림이다. 나는 이 그림을 설득력 있게 거부함으로써, 일반화된 정신 인과 문제에 답하고자 한다. 나는 실체 다원론이 이 문제에 대한 가장 적합한 해결책이라는 것을 보여주려고 할 것이다. 여기서 ‘실체 다원론’이라는 명칭으로 내가 염두에 두고 있는 것은 일종의 아리스토텔레스주의적 입장이다.

이 글의 내용은 최근의 신아리스토텔레스주의자들로부터 많은 영향을 받았다. (마이클 루, 카트라이트(N. Cartwright), 자웁스키(W. Jaworski), 로위(E. J. Lowe), 뎀포드(S. Mumford), 스템프(E. Stump), 프루스(A. Pruss) 등등.) 최근의 신아리스토텔레스주의는 아리스토텔레스 본인의 저작들에 대한 연구보다는 느슨하게 연결된 철학자들이 각자의 관심 분야에서 자생적으로 아리스토텔레스적인 생각들을 발전시키려는 운동으로 특징지어질 수 있다. 그래서 이 운동은 아직까지 통합적인 그림을 제시한다기 보다는 산발적으로 이루어지는 노력들을 통해 그러한 그림을 기대하게 하는 단계에 있다고 말해야 할 것이다. 이 글에서는 그러한 노력들 중 큰 두 줄기를 열어 보려고 시도했다. 이 글에서 전개되고 있는 주장에 가장 큰 영향력을 행사하고 있는 낸시 카트라이트는 최근의 가장 유명한 아리스토텔레스주의자들 가운데 한 사람으로, 현재 주류인 물리주의가 아닌 입장을 취하려 할 때 부정해야 할 것처럼 보이는 물리 영역의 인과적 폐쇄성을 그 누구보다도 설득력 있게 논파한 바 있다. 하지만 그녀의 주된 관심사는 과학 철학적 문제들이며, 전통적인 형이상학적 문제들에 대해서는 다소 무관심한 태도를 보인다. 반면에 전통적인 형이상학적 문제들로부터 출발하는 다른 아리스토텔레스주의자들은 좀 더 정교한 형이상학적 그림을 그려 주지만, 인과적 폐쇄성을 비판하는 데서는 카트라이트처럼 강력한 모습은 보여주지 못하고 있다는 인상을 준다. 이미 너무도 큰 친화력을 가지고 있는 양쪽의 노력을 결합한다면, 우리는 통합적인 그림을 제시하는 데에 한 발짝 더 다가가게 될 것이다.

각 장의 논의는 다음과 같이 강경한 물리주의로부터 점점 멀어지는 방

1) 베이커(Baker 1993)와 버지(Burge 1993)는 행위자로서 우리가 가지고 있는 그림의 설명적 중요성을 강조하는 것만으로 문제를 해결하려는 것처럼 보인다.

향으로 나아간다.

1장은 데이빗슨의 무법칙적 일원론에 대해 다룬다. 데이빗슨은 현재 가장 대중적인 위치를 차지하는 비환원적 물리주의라는 입장을 정립하는 데에 핵심적인 역할을 했다. 그는 정신적인 것이 물리적인 것으로 환원되지 않는다는 것을 설득력 있게 보여준다. 하지만 그는 정신적 사건 개별자와 물리적 사건 개별자가 동일하다고 주장한다. 나는 그가 사건 개별자들 사이의 동일성 기준을 제시하는 데 실패했으며, 유형 부수현상론 비판에 제대로 응답하는 데 실패했다는 두 가지 문제를 제기할 것이다. 전자는 개별자 동일성을 포기하고 수반 논제만을 남겨 놓으면 쉽게 해결할 수 있지만, 후자는 물리주의를 고수하는 한 대답하기가 쉽지 않은 것처럼 보인다.

2장에서는 개별자 동일성을 포기하고 수반 논제만을 유지하는 비환원적 물리주의들에 대해 다룬다. 이는 배제 문제에 대한 비환원적 물리주의자들의 답변을 검토함으로써 이루어지는데, 그들은 결국 정신적 속성의 인과적 효력을 제대로 설명하지 못하며, 따라서 유형 부수현상론 문제를 해결하지 못한다는 점을 보여줄 것이다.

3장에서는 물리주의를 거부하는 비주류적 입장들이 배제 문제를 해결할 때의 난관인 물리 영역의 인과적 폐쇄성 원리를 다룬다. 이 원리는 곧 근대 과학이 제시하는 것처럼 보이는 그림이며, 경험적 지지를 얻고 있다고 흔히 생각된다. 하지만 나는 카트라이트의 선구적인 작업들을 통해 이 원리가 가상에 불과하다는 점을 보여줄 것이다. 카트라이트가 보여주었듯이, 물리학의 근본 법칙들은 본질적으로 조건부적 성격을 가질 수밖에 없으며, 모델 혹은 모델에 충분히 가까운 상황에만 적용된다. 우리가 물리학의 이론을 적용하기 위해 모델을 구성하는 과정을 잘 살펴보면, 모델 구성 이전에 경험에 주어지는 세계는 오히려 온갖 다양한 인과력들이 다양한 효과들을 만들어내는 다원론적 세계라는 점이 드러난다. 따라서 경험적 지지를 얻는 것은 물리 영역의 인과적 폐쇄성이 아니라 오히려 개방성이다.

4장은 물리주의를 거부하는 이론들 가운데 어떤 입장이 가장 설득력 있는지에 대해 논한다. 현재 전형적인 속성 이원론자들은 정신적 속성의 근본성을 인정하면서도 실체에 관해서는 물리적 실체 일원론을 고수하는 경

향을 보이고 있다. 하지만 그들의 물리적 실체 일원론과 속성 이원론은 조화되기 어려운 입장들이다. 근본적 속성의 다수성을 받아들인다면, 실체의 다수성 역시 받아들이지 않을 도리가 없다는 점을 보여줄 것이다. 특히 창발론의 일종으로 이해될 수 있는 아리스토텔레스주의적인 실체 개념을 도입하면, 데카르트식 실체 이론을 비롯한 다른 실체 이론들이 겪는 문제들을 겪지 않고도 실체 다원론을 설득력 있게 제시할 수 있다. 이러한 아리스토텔레스주의적 실체 다원론은 행위자로서 우리가 가지고 있는 근본적인 신념과 더불어 다양한 거시적 인과력들을 현재 과학과 조화되게 자연 속에 확고하게 위치시킨다.

1. 정신적인 것의 무법칙성과 개별자 물리주의

정신적인 것과 물리적인 것 사이의 관계에 관한 데이빗슨의 입장은 ‘무법칙적 일원론’이라는 말로 요약된다. 무법칙적 일원론은 두 가지 큰 논제를 제시하고 있다. 우선 정신적인 것의 무법칙성은 정신적인 개념들이 물리적인 개념들로 환원되지 않는다고 주장한다. 그리고 일원론은 특정한 시공간 속에 있는 개별자(token)로서의 정신적 존재와 물리적 존재는 동일하다고 주장한다. 다시 말해서 데이빗슨의 입장은 정신적인 것과 물리적인 것 사이의 유형 동일성은 없지만, 개별자 동일성은 성립한다는 것이다.

이에 대해 나는 크게 두 가지 문제를 제기하고자 한다. 첫째, 데이빗슨은 그럴듯한 개별자 동일성 기준을 제시하지 못했고, 앞으로도 그러할 것이다. 정신적인 것의 무법칙성은 개별자 물리주의에도 영향을 미치기 때문이다. 데이빗슨이 개별자 물리주의를 주장하는 데에는 그가 심각한 검토 없이 받아들이는 ‘인과의 법칙적 성격 원칙’이 결정적인 역할을 하는데, 이 원칙이 포기된다면 개별자 물리주의를 주장할 필요도 사라진다. 둘째, 개별자 동일성 기준이 주어지든 주어지지 않든 데이빗슨에게 가해지는 부수현상론 비판은 여전히 해결하기가 어렵다. 데이빗슨과 그의 옹호자들은 부수현상론 비판에 제대로 대응하지 못했다. 이 비판은 개별자 물리주의를 포함하는 비환원적 물리주의 전반에 적용되는 비판이라는 점에서 더욱 심각하게 고려될 필요가 있다.

1.1. 정신적인 것의 무법칙성

정신적인 것의 무법칙성에 데이빗슨은 다음과 같이 말한다.

정신적인 것은 법칙적으로 환원 불가능하다: 정신적인 것과 물리적인 것을 연결해주는 **참된** 일반 진술들, 즉 법칙의 논리적 꼴을 갖춘 진술들이 있

을 것이다; 그러나 그 진술들은 (우리가 말하는 강한 의미에서) 법칙답지 (lawlike) 못하다. 설령 우리가 참된 비확률적(nonstochastic) 심물 일반 규칙을 만에 하나 발견했다고 하더라도, 우리는 그것을 대충 참인 것 이상의 것이라 믿을 만한 어떤 근거도 갖지 못할 것이다(데이빗슨 1970/2012c: 360).

그렇다면 일반 진술이 어떤 특징들을 가져야 법칙답다고 할 수 있을까? “물리 이론은 포괄적인 닫힌 체계를 제공한다고 약속한다. 그리고 그리하여 법칙에 맞는 어휘로 표현된 물리적 사건은 모두 표준에 맞는 유일한 기술을 반드시 갖도록 보증한다”(데이빗슨 1970/2012c: 371). 이런 보증이 이루어지려면, 물리적 기술을 하기 위해 다른 영역의 어휘들을 도입할 필요가 없어야 함은 물론이고, 사건 기술에 앞서 관련된 어휘들과 그 어휘들을 적용하는 규칙들이 잘 확립되어 있어야 할 것이다. 이와 같이 기술에 앞서 잘 확립된 체계와 연결되어 있는 것이 법칙다운 법칙, 엄격한 법칙이라 할 수 있다.²⁾

반면에 사건에 대한 정신적 기술의 경우, 우선 “정신적인 것에 영향을 미치지만 자신은 정신적인 것의 어떤 체계적인 부분도 아닌 것들이 너무도 많다”(같은 곳). 즉, 정신적인 것은 닫혀 있는 영역이 아니다. 게다가 정신적인 기술은 미리 확립된 체계를 가지고 시작할 수도 없다. 정신의 전체론적인 성격은 오히려 우리가 해석 대상이 최대한 합리적이라는 전제 하에서 상황에 맞춰 그때그때 체계 자체를 수정하도록 요구한다. 이 경우 사실 우리는 미리 주어진 잘 확립된 체계를 기대하지조차 않는다. 따라서 정신적인 어휘(명제 태도와 관련된 어휘)를 포함하는 일반적인 진술들은 엄격한 법칙이 아니며, 엄격한 법칙이 가지는 설명력과 예측력을 기대할 수는 없다. 이상의 설명에서 알 수 있듯이, 정신적인 것이 무법칙적인 이유는 그것의 전체론적인 성격에서 비롯하는 것이며, 정신적인 기술을 한다는 것은 번역 불확정성을 감수하면서 번역 행위를 하는 것과 같은 것이다.

2) “격을 갖춘 법칙이라면, 그 법칙을 적용하는 데 필요한 조건이 갖춰졌는지 아닌지를 우리가 미리 결정할 수 있도록 해줘야 한다”(데이빗슨 1974/2012d: 382)는 말도 같은 취지로 읽힌다.

그런데 데이빗슨은 엄격한 법칙이 오직 물리학에서만 찾아볼 수 있는 이상이라고 주장한다. 그가 보기에 현실의 물리학은 이러한 이상을 향해 수렴해 가고 있다(데이빗슨 1970/2012c: 365).³⁾ 반면에 심리학을 포함하는 다른 분과 학문들에서 찾아볼 수 있는 법칙적인 일반 진술들은 엄격하지 않은 법칙들이다. 이렇게 놓고 보면 데이빗슨은 단지 심리학이 물리학으로 환원되지 않는다고 주장하는 것처럼 보이는데, 그러면 심리학이 왜 꼭 물리학으로 환원될 수 있어야 환원 가능하다고 하는지, 다른 엄격하지 않은 법칙들로 이루어진 과학들로 환원되면 안 되는지에 관한 의문이 생길 수도 있을 것이다(Kim 1993a: 26). 하지만 합리성이라는 구성적 이상이 이론 체계에 제약을 가하는 경우는 대상에 명제 태도를 귀속시킬 때뿐이다. 따라서 데이빗슨이 말하고자 하는 바는 단지 심리학이 물리학으로 환원되지 않는다는 것이 아니라, 심리학이 대상에 대한 명제 태도 귀속이 불필요한 다른 학문들로 환원되지 않는다는 것이라고 보아야 할 것이다.

어쨌든 정신적인 것과 물리적인 것 사이의 유형 동일성은 이론간 환원을 전제한다는 것이 일반적인 생각이므로, 정신적인 것이 위와 같은 식으로 무법칙적이라면, 유형 동일성은 불가능하다는 결론이 나온다.

1.2. 개별자 동일성 기준

유형 동일성이 불가능하다면 개별자 동일성은 어떨까? 정신적인 것과 물리적인 것이 그토록 다른 질서 속에서 움직이는 것 같은데, 어떤 점에서 정신적 개별자와 물리적 개별자가 동일하다고 말할 수 있을까? 데이빗슨에게 정신적 개별자는 정신적 사건이며, 지향성을 드러내는 동사를 적어도 하나 이상 본질적으로 포함하는 형태의 기술에 의해 주어지는 사건이다. ‘S는 p라고 판단했다’가 그러한 부류에 속한다.⁴⁾ 이와 동일한지의 여부가 문제가 되는 물리적 사건은 편의상 그 시점에서의 ‘일군의 뉴런들이 발화

3) 물리학에서 정말로 엄격한 법칙을 기대할 수 있는지, 그리고 물리학이 그러한 이상을 추구하거나, 추구해야 하는지에 대해서는 의문을 제기할 여지가 많다.

4) 이러한 ‘정신적 행위’가 특정 시공간 상에 위치한다는 생각은 논란의 여지가 있다.

함' 같은 기술을 통해 주어진다고 해 보자. 데이빗슨에 따르면, 두 사건이 동일하다고 말할 수 있는 기준은 “그것들이 정확히 같은 원인과 결과를 가질 때 그리고 오직 그때만”(데이빗슨 1969/2012b: 308)이다.

이제 'S가 p라고 판단함'이라는 사건이 어떤 원인들과 결과들을 가질 것인지에 대해 생각해 보자. 정신적인 것은 달혀 있는 영역이 아니기 때문에, 우리가 흔히 정신적인 것과 무관하다고 생각하는 것들도 판단이라는 정신적 행위에 영향을 주는 원인일 수 있다. 그리고 그러한 부류의 원인들은 내가 의식할 수 있는 것들도 아닐 경우가 많다. 우리가 특정 판단을 하게 된 '이유'라고 부를만한 것들에만 주의를 집중해 보자. 데이빗슨에 따르면 이유도 원인이다. 누군가가 S에게 '당신은 왜(어떤 이유에서) p라고 판단했습니까?'라고 묻는다면, 그는 이러저러한 이유에서 그렇게 판단했다고 간단하게 답할 것이다. 하지만 그때 그가 거론한 이유는 그 판단의 완전한 원인과는 거리가 멀다. 그것은 그저 그 판단이 이루어지게 한 원인들 중 하나에 불과하다.⁵⁾ 어떤 정신적 사건과 물리적 사건이 동일한지의 여부를 가리기 위해서 필요한 것은 관련된 완전한 인과적 요인들의 목록이다. 그래도 어쨌든 판단의 원인이 된 이유들의 목록만이라도 완전하게 가지기를 원한다고 해 보자. 정신적인 것의 무법칙성의 배후에 있는 번역 불확정성은 사실 이 단계에서 이미 문제를 일으킨다. S가 p라는 판단을 내리게 된 이유들, 즉 그의 믿음과 욕구들에 대한 결정적인 목록 같은 것은 없을 것이기 때문이다. 그럼에도 불구하고 어떻게든 이유들의 목록이 주어졌다고 해 보자. 데이빗슨은 다음과 같이 말한다.

바람과 믿음이 한 행위를 올바르게 설명하기 위해서는 그것들이 올바른 방식으로, 다시 말해 아마도 합리성의 기준을 만족시키는 추리의 사슬이나 과정을 거쳐서, 그 행위의 원인이 되어야 한다. 그 올바른 종류의 인과적 과정을 구별할 수 있으려면 무엇보다도, 상반되는 증거와 상반되는 바람들로부터 어떻게 하나의 결정이 내려지는지가 밝혀져야 한다. 나는 그것을 밝혀내

5) “이유에 의한 설명은 인과적 요인들이 복잡하게 있으면 이중에서 하나를 골라내는 데 만족하고는, 그 복잡한 것과 직접 맞부딪치는 것을 피해버린다”(데이빗슨 1974/2012d: 383).

는 것이 도대체 가능한지 의심스럽다(데이빗슨 1974/2012d: 382).

실천적 판단에서 더욱 두드러지는 것처럼 보이기는 하지만, 이론적 판단에서도 S가 p라고 판단할 때, 그는 p와는 다르게 혹은 반대로 판단할 이유들을 가지고 있는 경우는 매우 흔하다. 그럴 때 우리가 어떻게 판단을 해야 하는지 말해 줄 수 있는 규칙은 없거나, 있다고 하더라도 지나치게 일반적이어서 특수한 상황에 적용을 하려면 다시 판단이 필요한 것들 뿐이다. 같은 취지에서, 위의 인용문은 그러한 판단 상황에서 S가 가지고 있는 이유들로부터 그가 실제로 어떤 판단을 내리게 되는지를 알려주는 일반적으로 적용 가능한 의사 결정 이론이 없다고 말하고 있다. S가 가진 이유들이 그의 실제 판단과 어떻게 연결되는지 제대로 이해하려면, 우리는 결국 그 사람의 개별적 성격, 즉 그가 가진 정신적 성향들의 체계를 필요로 하는데, 이는 물리적으로 기술된 원인과 결과를 연결시킬 때에는 불필요한 것이다.

정리하자면 문제는 다음과 같다. 한 정신적 사건이 다른 물리적 사건과 동일한지 알려면, 양자가 정확히 동일한 원인들과 결과들을 가지는지 알아야 한다. 그런데 한 정신적 사건의 원인들은 다른 많은 정신적 사건들로 이루어져 있고, 물리적 사건들의 원인 역시 다른 많은 물리적 사건들로 이루어져 있다. 따라서 서로 다른 영역에 속하는 두 사건이 동일한지 알기 위해서는 각 사건에 연결된 인과들의 계열이 서로 동일한지 알아야 하는데, 이것이 원리적으로라도 가능하려면, 적어도 두 영역의 인과 계열들 사이에 동형적 구조는 있어야 한다고 생각해야 할 것이다. 그래야 양쪽을 대응시키려는 시도라도 해 볼 수 있기 때문이다. 하지만 정신적인 것의 무법칙성으로부터 우리가 알 수 있는 것은 그러한 동형적 구조를 기대할 수 없다는 것이다.⁶⁾

아니면 위와는 달리 개별자 동일성의 기준으로서 데이빗슨이 다소 의심하면서도 검토하고 있는 또 다른 기준인 같은 시공간상에 위치함은 어떨까? 하지만 이 기준은 더욱 문제적이다. 우선 선결 문제 요구의 오류를

6) 동형적 구조의 필요성으로부터의 논증은 차일드(Child 1994: 80-89), 참조.

범하지 않고 정신적 사건이 일어나는 시공간을 어떻게 정할 수 있는지가 의문이다. 지향성을 드러내는 동사를 본질적으로 포함하는 기술에 의해 주어지는 모든 사건들이 일어나는 공간적 위치에 대해 우리가 말할 수 있는가? 나는 우리가 일반적으로 그런 것을 말하기 위한 관행을 가지고 있는지 의심스럽다. 데이빗슨은 정신적 사건이 일어났을 때 그 사건을 이루는 변화가 일어난 사람이 있는 장소가 곧 정신적 사건이 일어난 공간적 위치라고 여긴다(데이빗슨 1969/2012b: 304). 정신적 사건의 시간적 위치에 대해 말하는 것은 일상적으로나 전통 철학적으로나 좀 더 지지를 받는 듯한 인상을 가지고 있으나, 그 경우에도 정신적 사건이 일반적으로 물리적 사건들처럼 시간적으로 개별화되는지는 다소 의문이다. 어쨌든 어떤 식으로라도 정신적 사건이 일어나는 시공간상의 위치가 주어질 수 있다고 해보자. 그런데 그 시공간상의 위치에서는 우리가 ‘물리적’이라 부르는 대상들의 범위가 확정되었다고 하더라도, 많은 사건들이 일어날 것 같다. 그 많은 사건들 중 어떤 것이 문제가 되는 정신적 사건에 대응하는가? 거기서 일어나는 모든 사건들인가? 아니면 일부 사건들인가?

이와 관련된 또 한 가지 심각한 문제는 데이빗슨 자신이 제기하고 있는 것인데, 정확히 같은 시공간적 위치에 있지만, 서로 다른 두 사건이 발생하고 있다고 볼 수 있는 경우들이 있다는 것이다. “예를 들어 만약 한 쇠공이 몇 분 동안 따뜻해지면서 같은 시간 동안 35도 회전한다면 이것들이 같은 사건이라고 말해야 하는가? 그렇지 않은 것 같다”(데이빗슨 1969/2012b: 307-8)고 그 스스로 말하고 있다. 물론 그는 이어서 이 사례에 대한 다른 해석 가능성도 검토한다. 쇠공의 따뜻해짐과 회전함이라는 두 사건이 있는 것이 아니라, 실제로는 그 공을 이루는 분자들의 운동이라는 한 사건만이 있다는 것이다. 하지만 이러한 해석은 미봉책에 불과하다. 분자 운동 수준에서도 같은 시공간 상에서 두 사건이 일어날 수 있기 때문이다.

이런저런 기준들 없이 정신적 개별자와 물리적 개별자가 그냥 동일하다고 말하고 싶을 수도 있다. 이러한 고유한(Sui generis) 동일성은 별로 흥미로운 것이 못 되기도 하지만, 그것과 양립 불가능한 평행론을 주장하는

것과 다를 바가 없다.⁷⁾ 게다가 내가 보기에 ‘동일성’이라는 말을 이런 식으로 사용하는 것은 이상한 방법이다. 일상적으로 A와 B가 동일하다고 말할 때, 우리는 양자가 어떤 차이에도 불구하고 그 차이를 무시할만한 중요한 어떤 점에서 동일하다고 말하고 있는 것이기 때문이다. 데이빗슨 자신도 사건 동일성 기준이 제시되어야 한다는 점에는 동의하고 있다(데이빗슨 1969/2012b: 288). 문제는 그가 제시하고 있는 동일성 기준들이 받아들여질 수 없는 것들이라는 점이다.⁸⁾

지금까지의 논의에 비추어 볼 때, 무법칙적 일원론에서 ‘일원론’ 부분은 거부할만한 이유가 있다. 중요한 점은 데이빗슨의 철학에서 ‘일원론’ 부분은 제거할 수 없을 정도로 중요한 부분이 아니라는 것이다. 그가 개별자 동일성을 주장하는 이유를 간략히 제시하자면 다음과 같다.

(전제1) 정신적 사건과 물리적 사건은 인과적으로 연결된다.

(전제2) 단칭 인과 진술은 모두 어떤 엄격한 법칙에 의해 뒷받침된다.

(전제3) 엄격한 심물 법칙은 없다.

(결론) 따라서 정신적 사건과 물리적 사건은 동일하다.

다시 말해서, 정신적 사건이 물리적 사건과 인과 작용을 하는 것은 분명하고, 인과 작용이 있기 위해서는 그것을 뒷받침하는 엄격한 법칙이 있어야 하는데, 엄격한 심물 법칙은 없으므로, 문제가 되는 정신적 사건에 대응되는 기술을 가지는, 즉 같은 원인과 결과를 가지는 물리적 사건이 있어

7) 퍼트남(Putnam 1999: 31-32), 참조.

8) 데이빗슨은 초기에 사건들 사이의 동일성 기준으로 원인과 결과의 동일성을 더욱 지지했으나, 나중에는 그가 초기에 의심한 시공간적 위치상의 동일성만을 배타적으로 지지하는 쪽으로 돌아섰다(Davidson 2001: 309-10). 하지만 여기서 그는 사건들 사이의 동일성 기준을 물리적 대상들 사이의 동일성 기준과 같은 것으로 만드는 콰인의 견해에 공감을 표하면서도, 그러면 우리가 어떻게 사건과 대상을 확실하게 구분할 수 있을지 고민하고 있다. 결국 데이빗슨은 사건들 사이의 동일성 기준에 대한 결정적인 답변을 내놓지 못한 것으로 보인다.

야 그것을 포섭하는 엄격한 법칙의 뒷받침을 받을 수 있지 않겠느냐는 것이다. 여기서 전제1은 정신 인과에 관한 의견을 피력하는 거의 모든 논자들이 의심하지 않는 가정이다. 또, 전제3에 해당하는 정신적인 것의 무법칙성은 비록 전제1에 비하면 의심하는 사람들이 많기는 하지만, 그래도 비교적 광범위하게 받아들여지고 있는 가정이라 할 수 있다. 반면에 전제2는 그가 특별한 근거 없이 그냥 받아들이고 있는 당시의 주류적 견해에 불과하며, 그만큼 많은 사람들의 의심을 사고 있는 가정이다. 따라서 전제2가 면밀하게 검토되어야 한다.

1.3. 일반주의 대 단칭주의

데이빗슨이 ‘인과의 법칙적 성격 원칙’이라 부르는 전제에 대해 검토해보도록 하자. 그것은 다음과 같은 것을 의미한다.

만약 한 사건이 다른 사건의 원인이 되면, 적절히 기술되면 이 사건들을 사례로 갖는 어떤 엄격한 법칙이 있다. 그러나 그 법칙이나 그 적절한 기술을 알지 못해도 단칭 인과 관계를 아는 것은 가능하다(그리고 그것은 흔한 현상이다). 앎은 이유를 필요로 한다. 그런데 이유는 거친 이질법칙적 일반 규칙의 형태로 손에 주어진다. 이런 규칙들이 법칙답다는 것은 끝없이 정교해질 수 있다는 의미에서가 아니라, 한 사례로부터 우리가 선례를 따르는 다른 사례들을 기대하는 것이 합리적이라는 의미에서다(데이빗슨 1970/2012c: 372).

여기서 데이빗슨이 가지고 있는 그림은 다음과 같은 것이다. 우리는 단칭 인과 진술들을 알고 있다. 단칭 인과 진술들은 원칙적으로는 엄격한 법칙에 의해 정당화되어야 하는데, 현실에서 우리는 어떤 엄격한 법칙이 우리가 아는 그 단칭 인과 진술들을 사례로 갖는지 알지 못하는 경우가 많다. 현실에서 단칭 인과 진술들은 많은 경우 오히려 거친 이질법칙적 일반 규칙, 혹은 느슨한 법칙에 의해 정당화된다. “성냥을 그으면 불이 붙는다”

와 같은 진술이 그러한 거친 이질법칙적 일반 규칙에 해당한다. 이런 일반 규칙들은 엄격한 법칙이 가지는 예외 없음이라는 성격을 가지지 못한다. 성냥을 굿더라도 불이 붙지 않을 조건들은 많이 있다. 그래서 느슨한 법칙들은 ‘다른 조건이 동일하다면(ceteris paribus)’이라는 단서가 붙는 조건부 법칙일 수밖에 없다. 데이빗슨은 느슨한 법칙에서 조건절을 명시화함으로써 제거하여 엄격한 법칙으로 만들 수 있다고 믿지 않는다.

그런데 문제는 여기서부터 시작된다. 단칭 인과 진술들이 이러한 조건부 법칙에 의해 정당화되는 것이 참이라고 해 보자. 그렇다고 하더라도 단칭 인과 진술들은 궁극적으로는 엄격한 법칙에 의해 정당화되어야 하므로, 그리고 조건부 법칙에서 조건절은 제거될 수 없으므로, 이제 조건부 법칙은 엄격한 법칙에 의해 정당화되어야 할 것이다. 그렇다면 “성냥을 그으면 불이 붙는다”는 느슨한 법칙을 정당화하는 엄격한 법칙은 무엇인가? 엄밀히 말하자면 그 누구도 알지 못한다. 앞서 말하였듯이 엄격한 법칙은 오직 완성된 물리학만이 제공할 수 있는 것인데, 현재 우리가 알고 있는 물리학은 완성된 물리학이 아니기 때문이다. 현재의 물리학만을 고려하더라도 조건부 법칙을 정당화하는 물리 법칙이 구체적으로 어떤 것인지 알 수 있는 경우는 일반적이지 않다(3장 참조). 아무도 모르는 근거에 의해 어떤 진술들이 정당화된다는 말은 확실히 이상하게 들린다. 데이빗슨 자신도 어느 정도 이러한 문제를 의식하고 있었던 것인지 다음과 말을 하고 있다.

그리고 내 생각에 우리는 매우 자주, 어떤 적절한 인과 법칙이 있다고 믿을 이유를 갖기만 하면 그 법칙이 뭔지 모르더라도 그것을 어떤 단칭 인과 진술을 받아들이는 정당한 근거로 여긴다. “만약 당신이 알맞게 준비된 표면에, 충분히 강하게, 제대로 만들어진 성냥을 그으면, 다른 조건들이 괜찮다면 불이 붙을 것이다” 같은 일반 규칙들이 나름의 중요성을 갖는다면 그것은 우리가 결국 그것들을 경향성을 넘어서, 어떤 예외도 없는 것으로 만들 것이라고 기대할 수 있기 때문이 아니라, 오히려 우리가 설명하기를 원하는 사건을 감싸 안는 제대로 된 어떤 인과 법칙이 있다고 믿을 만한 우리의 증거 중 많은 부분을 그것들이 요약해주기 때문이다(데이빗슨 1967/2012a: 279-80).

즉, 우리는 정확히 어떤 엄격한 법칙이 문제의 단칭 인과 진술을 정당화하는지 알지 못하더라도, 모종의 엄격한 법칙이 배후에 있다고 믿을 이유를 가진다면 그 단칭 인과 진술을 엄격한 법칙에 의해 정당화된 것으로 간주할 수 있다는 것이다. 그리고 데이빗슨은 느슨한 법칙이 단칭 인과 진술의 배후에 모종의 엄격한 법칙이 있다고 믿을 바로 그러한 이유를 제공한다고 주장하고 있다.

하지만 이 대목에서 그가 왜 이런 주장을 하는지 이해하기가 쉽지 않다. 특히 위의 인용문에서 느슨한 법칙이 “우리가 설명하기를 원하는 사건을 감싸 안는 제대로 된 어떤 인과 법칙이 있다고 믿을 만한 우리의 증거 중 많은 부분을 그것들이 요약해주기 때문”이라는 구절이 해석하기 어렵다. 일단 논의를 위해 단칭 인과 진술이 그것을 사례로 가지는 모종의 엄격한 법칙이 있는 한에서만 정당화된다는 주장이 참이고, 그렇게 믿을 만한 근거들이 있다고 해 보자. 그렇다고 하더라도 일반적인 느슨한 법칙이 어떻게 그러한 근거 내지는 근거들을 요약한 것으로서 썸해질 수 있는지가 의문이다. “성냥을 그으면 불이 붙는다”와 같은 일반 규칙 진술이 어떤 점에서 그것을 정당화하는 모종의 엄격한 법칙의 존재를 믿을 만한 근거를 제공하는가?

또, 일반 규칙 진술이 모종의 엄격한 법칙의 존재를 믿을 만한 근거들의 ‘요약’이라는 설명도 해석하기가 어렵다. 일반 규칙 진술이 무언가의 요약으로 읽힐 수 있다면, 가장 먼저 떠오르는 생각은 그것이 일군의 단칭 인과 진술들의 요약이라는 생각일 것이다. 하지만 이러한 해석은 곧바로 저항에 부딪힌다. 일반 규칙 진술이 정말로 그것에 앞서 주어진 단칭 인과 진술들의 요약이라면, 단칭 인과 진술들이 일반 규칙 진술에 의해 정당화된다는 생각이 받아들이기 어려운 것이 되기 때문이다. 그 경우 오히려 일반 규칙 진술이 요약된 단칭 인과 진술들에 의해 정당화된다고 보는 편이 옳을 것이다. 하지만 이는 데이빗슨이 가지고 있는 그림과 상반된다. 따라서 일반 규칙 진술이 요약하고 있는 것은 단칭 인과 진술들이 아닌 어떤 것들이라 보아야 할 것이다.

어쩌면 데이빗슨은 정말로 흠처럼 일반 규칙 진술을 특정한 두 사건이 연속적으로 일어나는 경우들을 요약한 것으로 보고 있는 것일 수도 있지만, 두 사건이 연속적으로 일어나는 것이 엄격한 법칙의 존재를 믿을 만한 근거를 제공한다고 볼 수는 없다. 그것은 잘 해야 느슨한 법칙의 존재를 믿기 위한 근거를 제공할 것이다.

위와 같은 해석상의 난점들 때문에 우리는 데이빗슨이 ‘인과의 법칙적 성격 원칙’을 받아들이는 근거가 정확히 무엇인지 그의 글만을 보아서 알기가 힘들다. 한 가지 추측을 해 보자면, 그는 험펠 등등이 이 문제에 관해 이미 많은 것을 말했으며, 그것이 표준적으로 수용할 만한 내용이라고 생각했기 때문에 더 이상 자세하게 논할 필요성을 느끼지 못했던 것일 수도 있다. 실제로 이 원칙은 험펠식 연역법칙적(DN) 모델을 인과적 설명에 적용한 것과 거의 같은 내용을 가지고 있다.⁹⁾ 그렇다면 험펠이 연역법칙적 모델에 관해 가졌던 생각으로부터 데이빗슨이 ‘인과의 법칙적 성격 원칙’을 지지하는 이유에 대해서도 추론해 볼 수 있을 것이다.

우드워드에 따르면 험펠이 연역법칙적 모델을 제안한 데에는 크게 두 가지 동기가 작용하고 있었다(Woodward 2014: 2.4). 우선 첫째는 그가 과학적 설명과, 예측, 이해라는 세 가지 개념을 하나로 묶어서 취급했다는 점이다. 연역법칙적 모델에서 전제들로부터 결론을 연역적으로 도출한다는 것은 설명하고자 하는 현상이 법칙적으로 예측 가능했다는 점을 보여주는 것이며, 이것이 곧 현상에 대한 설명과 이해를 제공하는 것이라고 험펠은 생각했다.

그 다음 둘째는 인과적 진술이 과학적 설명에서 하는 역할에 대해 그가 가지고 있었던 생각과 관련되어 있는데, 우리에게는 이 두 번째 동기가 중요하다. 험펠은 인과 개념을 원초적인 것으로 받아들이기를 거부했다. 즉, 그는 ‘X는 Y의 원인이므로, X는 Y를 설명한다’와 같은 식의 생각을 받아들이지 않았다. 그 이유는 그가 흠에서 비롯된 전통에 속해 있기 때문이었다. 잘 알려져 있듯이, 흠은 법칙이나 인과 같은 개념들을 규칙성으로 대

9) 특히 험펠(2010: 112-113) 참조. 데이빗슨은 한편으로는 단칭 인과 진술이 그것을 사례로 갖는 법칙 없이도 알려지는 것처럼 보이는 전이론적 직관에 공감을 하면서도, 다른 한편으로는 험펠보다 엄격한 법칙의 역할에 대해 강조하는 경향이 있다.

체했다. 험펠의 연역법칙적 모델에서도 ‘법칙’은 그저 규칙성을 의미하는 것이다. 흄적인 전통에서 법칙이나 인과 같은 일련의 개념들은 필연성과 같은 양상 개념과 밀접한 관련을 맺고 있는 것으로 이해되었고, 경험적으로 정당화될 수 없는 것으로 여겨졌다. 흄의 이론 중 인상과 관념에 대한 학설을 비롯한 다른 많은 부분들을 수용하지 않았음에도 불구하고, 논리 경험주의자들은 양상 개념에 대한 거부감만큼은 그대로 이어받았다. 그리고 데이빗슨에게 지대한 영향을 준 콰인은 논리 경험주의에 대한 가장 강력한 비판자들 중 한 사람이었지만 그 역시 양상 개념에 대해서는 우호적이지 않았다. 이러한 전통을 물려받았기 때문에 데이빗슨은 ‘인과의 법칙적 성격 원칙’을 심각하게 문제삼지 않고 받아들였을 것이다.

그렇다면 흄적인 전통에서 비롯된 일반주의(Generalist)적 인과 이론, 즉 단칭 인과 진술에 앞서 그것과 독립적으로 이해된 규칙성에 관한 일반적 진술인 법칙이 단칭 인과 진술을 사례로 가짐으로써 정당화한다는 입장은 얼마나 설득력이 있는 것일까? 단칭 인과 진술이 그것을 사례로 가지는 법칙 없이도 그 자체로 설명력을 가지는 것처럼 보인다는 점은 법칙이 인과적 설명의 필요 조건이라는 점을 의심하게 만든다. 이러한 의심에 대해 데이빗슨과 험펠을 비롯한 일반주의자들은 겉으로 보이는 것과는 달리 완전한 이상적 설명에서는 법칙이 설명항에 나타난다거나 나타나야 한다고 응수할 것이다. 이러한 식의 접근법의 대가는 물론 실제 설명 관행과 이상적 설명 관행 사이의 심각한 괴리일 텐데, 일반주의자들은 그래도 지불할 만한 대가라고 생각할 것이다.

위와는 달리 법칙이 인과적 설명의 충분 조건조차 아니라는 점을 지적하는 비판들도 있다. 대표적인 것은 연역법칙적 이론이 설명의 비대칭성을 설명할 수 없다는 다음과 같은 반례이다. 깃대의 높이와 태양이 지면과 이루는 각도 그리고 광학 법칙으로부터 깃대의 그림자의 길이를 도출할 수 있다고 해 보자. 이러한 도출은 연역법칙적 모델이 제시하는 설명의 기준을 충족시킨다. 하지만 반대로 깃대의 그림자의 길이와 태양이 지면과 이루는 각도 그리고 광학 법칙으로부터 깃대의 높이를 도출하는 것은 연역법칙적 모델에 잘 들어맞지만 설명으로 간주되지 않는다. 인과 관계로 생

각하더라도 햇빛과 깃대가 그림자가 진 원인이라는 생각은 자연스럽지만, 그림자가 진 것이 깃대의 원인이라는 생각은 부자연스럽다. 이러한 예들은 얼마든지 만들어낼 수 있다. 총알의 발사와 총구의 화염 발생이라는 한 쌍의 사건에서, 총알의 발사로 총구의 화염 발생을 설명하는 것은 가능하지만, 그 역은 아니다. 일반화시키자면 법칙이 두 사건 사이의 단순한 규칙성을 서술하는 쌍조건문인 한, 그것은 원인과 결과 사이의 비대칭성을 설명하지 못한다. 이밖에도 일반주의적 인과 이론은 이와 직접 혹은 간접적으로 관련된 여러 가지 문제들을 해결하지 못한다고 알려져 있다.¹⁰⁾

다른 유력한 대안이 없다면 일반주의의 입장에서 제기된 문제들을 해결하고자 하는 노력들을 좀 더 깊게 다룰 필요가 있을 것이다. 하지만 이제까지의 논의를 통해 당장 쉽게 생각해 볼 수 있는 대안이 있다. 그것은 단칭 인과 진술이 일반 진술보다 적어도 덜 근본적이지는 않다는 단칭주의(Singularist)적 인과 이론이다. 한때 흙적인 전통이 분석 철학 내에서 주류적인 경향이었다고는 하나, 그와는 대립되는 상대적으로 소수인 흐름들도 꾸준히 있어 왔다. 대표적인 예로 엘리자베스 앤스콤(Anscombe 1971/1993)을 들 수 있다. 앤스콤은 이 고전적인 논문에서 인과를 필연적 연결이나 예외 없는 규칙성의 사례로 분석하는 광범위한 경향들을 비판하고 있다. 앤스콤은 우리가 자주 단칭 인과 진술의 배후에 그것을 사례로 가지는 예외 없는 일반 진술이 있는지의 여부를 아는지와 무관하게 단칭 인과를 안다는 사실에 주목한다(Anscombe 1971/1993: 91). 앞에서 보았듯이, 데이빗슨은 단칭 인과 진술이 법칙과 무관하게 알려지는 것처럼 보이는 현상에 직면하여 어떻게든 그 배후에서 단칭 인과 진술을 근거짓는 모종의 법칙을 찾으려 했던 반면, 앤스콤은 같은 상황에서 현상을 액면 그대로 받아들이고, 그것이 단칭 인과 진술을 일반주의적으로 분석할 수 없다는 것을 보여주는 징표라고 여긴다. 그녀는 그러한 분석이 인과 관계의 가장 핵심적인 측면을 포착하지 못한다고 본다.

10) 이 글에서 다루고 있지는 않지만, 단칭 인과 진술을 반사실적 조건문이나 통계적 진술로 근거지으려는 이론들 역시 연역법칙적 모델과 유사한 문제점들을 노출시키고 있다. 무어(Moore 2009: 508-10)는 이 문제들을 일곱 가지 유형으로 분류한다.

인과 관계는 그것의 원인들로부터 결과의 이끌어내짐이라는 데 있다. 이것이 다양한 종류들의 인과 관계들의 핵심이고, 공통된 특징이다. 결과들은 그것들의 원인들로부터 이끌어내지고, 일어나고, 나온다. 예를 들어, 모든 사람은 물리적인 부모자식관계가 인과적 관계라는 것을 받아들일 것이다. 여기서 그 이끌어내짐은 분열에 의한 물질적인 것이다. 이제 필연성이나 일반성이라는 점에서의 분석은 결과의 이 이끌어내짐에 대해 우리에게 말해주지 않는다. 오히려 그것[분석]은 저것[이끌어내짐]에 대해 잊어버린다. 필연성은 자연 법칙의 그것일 것이기 때문에, 그것을 통해 우리는 원인에 대한 지식으로부터 결과에 대한 지식을 이끌어내거나 그 반대로 할 수 있을 것이지만, 그것은 결과의 원천으로서의 원인을 우리에게 보여주지 않는다. 그렇다면 인과는 필연성과 동일시되어서는 안 된다(Anscombe 1971/1993: 92).

다시 말해서 원인 개념에 있는 “결과의 원천”이라는 성격은 단칭 인과를 단순히 규칙성으로 이해된 법칙의 사례로 볼 수 없게 만드는 문제적인 요소이다. 일반주의가 설명의 비대칭성을 제대로 설명하지 못한다는 심각한 문제점은 인과 개념에 있는 이러한 성격을 도외시하기 때문이라는 점을 쉽게 알 수 있다. 반면에 법칙에 대한 참조 없이도 어떤 단칭 인과 진술이 참인지 알 수 있다고 보는 앤스콤과 같은 입장에서는 인과 관계의 비대칭성을 설명해야 한다는 문제는 애초에 발생하지도 않는다.

우리가 개별적인 인과 관계를 직접 관찰할 수 없기 때문에 단칭주의는 성립할 수 없다는 비판은 흄적인 전통 쪽에서 제기되는 가장 대표적인 것이었다. 여기서 개별적인 인과 관계를 직접 관찰할 수 없다는 것이 정확히 어떻게 단칭주의에 대한 비판으로 작용하는지 생각해 볼 필요가 있다. 이러한 비판이 암묵적으로 전제하고 있는 것은 토대주의적인 정당화의 그림이다.¹¹⁾ 즉, 모든 경험적 지식은 직접 관찰할 수 있거나 직접적인 관찰을 통해 형성된 지식에 토대를 두고 있다는 그림이다. 또, 이러한 비판이 성

11) 일반적으로 데이빗슨은 콰인의 영향으로 이러한 토대주의적인 정당화의 그림을 거부하고 정합주의적인 그림을 채택하는 것으로 알려져 있다. 하지만 ‘인과의 법칙적 성격 원칙’ 배후에 있는 것은 바로 이러한 토대주의적인 그림인 것처럼 보이기 때문에 데이빗슨에 대한 해석은 더욱 난항을 겪게 된다.

립하기 위해서는 토대 개념이 환원을 함축하도록 강하게 해석되어야 한다. 그렇다면 근거지어지는 지식들이 가지는 모든 요소들은 근거짓는 지식들로부터 나와야 하는 것으로 이해될 것이다. 이와 같은 전제들을 받아들이는 한에서만, 개별적인 인과 관계에서 원인이 가지는 결과의 원천이라는 성격은 관찰될 수도 없고, 관찰된 사실들에 근거지어질 수도 없는 추가적인 요소이므로 정당화될 수 없다는 결론이 나올 것이다. 물론 이러한 방향으로 나아가기 시작하면 상식과의 괴리는 피할 수 없는 귀결일 텐데, 잘 알려져 있듯이 흄 자신은 회의주의적 결론이 철학적인 관점에서는 옳바르다고 판단했다.

앤스콤은 개별적인 인과 관계를 직접 관찰할 수 없다는 주장에 강하게 반발한다. 즉, 토대주의적인 정당화의 그림을 받아들인다고 하더라도 단칭 인과에 관한 지식은 정당화된다는 것이다. 흄이 가지고 있는 그림에서, 내가 발로 찬 것이 원인이 되어 공이 날아갔다는 식의 기술은 문자 그대로 받아들여지지 않는다. 그것은 내가 발로 참이라는 사건에 공이 날아감이라는 사건이 이어져서 일어났다는 식으로 재기술되어야 문자 그대로 받아들여질 수 있다. 이 대목에서 앤스콤은 흄이 공의 날아감이라는 ‘움직임’은 관찰될 수 있는 항목으로 여기고 있다는 점에 주목한다. 하지만 그렇다면 적어도 어떤 개별적 인과 관계는 직접 관찰할 수 있다고 보아야 한다는 것이 그녀의 주장이다. 이는 다음의 인용문에 잘 드러나 있다.

‘원인’이라는 단어 자체는 고도로 일반적이다. 어떤 사람이 그가 원인 개념을 가지고 있다는 것을 어떻게 보여주는가? 그의 어휘 속에 그와 같은 단어를 가지고 있음으로써만, 이라고 우리는 말하기를 바랄지 모른다. 만약 그렇다면 그 개념의 명백한 소유는 언어에서 다른 많은 것들의 속달을 전제한다. ‘원인’이라는 단어는 많은 인과적 개념들이 이미 표상되고 있는 언어에 더해질 수 있다는 뜻이다. 작은 선별: **문지르다, 밀다, 적시다, 옮기다, 먹다, 태우다, 넘어뜨리다, 막다, 으깨다**, (예를 들면, 소음, 종이배를) **만들다, 다치게 하다**. 그러나 만약 우리가 어떤 특수한 인과적 개념들도 표상되지 않는 언어들에 상상하고자 한다면, 그와 같은 언어에서 단어에 대한 어떤 기술도 원인을 의미하는 것으로서 그것을 제시할 수 없을 것이다. 그것

은 심지어 자연종을 위한 단어들도, ‘신체’, ‘바람’, 혹은 ‘불’에 해당하는 단어들도 포함하지 않을 것이다. 왜냐하면 특수한 인과적 동사들을 사용하는 것을 배우는 것은 이러한 그리고 다른 많은 명사 어구들(substantives)에 대응하는 개념들을 적용하는 것을 배우는 것의 일부분이기 때문이다. 우리가 사람들을 이름으로 부르거나 고양이가 탁자 위에 있다는 것을 봄으로부터 보고하는 것을 배우는 것과 마찬가지로 확실하게, 우리는 또한 누군가가 우유를 마셔버렸다거나 개가 웃기는 소음을 낸다거나 사물들이 그것들을 자르거나 부수는 우리가 보는 어떤 것에 의해서든 잘리거나 부서진다는 것을 관찰함으로부터 보고하는 것을 배운다(Anscombe 1971/1993: 93).

흄주의자들이 허용하는 개별적인 관찰 진술은 여러 동사들을 사용하고 있다. 그러한 동사들은 이미 인과적 의미를 포함하고 있으며, 그것들 없이는 개별적인 관찰 진술도 성립할 수 없을 것이다. 인과적 개념들은 흄주의자들이 생각하는 것보다 우리의 언어에 훨씬 더 만연해 있다. 그리고 바로 그렇기 때문에 우리는 ‘인과’라는 추상적 개념을 습득할 수 있는 것이다. 이처럼 많은, 어쩌면 거의 모든 언어 표현들이 이미 인과-적재적이고 우리가 그러한 언어들을 통해 지각된 현실을 기술하며 살아가고 있다면, 인과적 개념과 비인과적 개념을 날카롭게 나누고 인과적 개념들을 비인과적 개념에 근거지으려는 시도는 동기를 잃게 될 것이다.

물론 단칭주의에 앤스콤과 같은 비환원주의적 입장만 있을 수 있는 것은 아니다. 단칭 인과 진술이 관련된 법칙보다 덜 기본적인지는 않다는¹²⁾ 입장을 고수하면서도 개별적인 인과 관계 자체는 비인과적인 관계로 환원

12) 무어(Moore 2009: 504)는 여기서의 ‘기본적’(혹은 원초적)이라는 말의 의미를 세 가지로 나누어 고찰하고 있다. 첫째로, 분석적으로 기본적이라는 것은 인과적 개념이 비인과적 개념으로 분석될 수 없다는 뜻이다. 둘째로, 인식적으로 기본적이라는 것은 개별적인 인과 관계가 법칙에 대한 참조 없이, 관찰 등을 통해 직접 알려질 수 있다는 뜻이다. 셋째로, 형이상학적으로 기본적이라는 것은 인과 관계가 다른 속성이나 관계로 환원될 수 없다는 뜻이다. 하지만 흄과 앤스콤의 경우에서 알 수 있듯이, 이 세 가지 의미는 함께 가는 경우가 많다. 앤스콤의 논증은 인과 관계가 인식적으로 기본적이므로 분석 불가능하고, 분석 불가능하므로 다른 것으로 환원될 수도 없다는 점을 보이는 것으로 해석할 수 있다. 어쨌든 무어가 보기에 이 세 가지 의미 중에서 환원주의와 비환원주의를 나누는 기준이 되는 것은 형이상학적 의미이다.

될 수 있다는 입장도 있을 수 있다.¹³⁾ 여기서 비인과적인 관계는 단순한 시공간적 연속 관계, 모종의 물리적인 과정, 반사실적 의존 관계, 확률적 사실 등등 다양한 후보들이 있을 수 있다. 또, 단칭 인과와 법칙 사이의 관계에 있어서도, 단칭 인과가 법칙보다 더 기본적인이라는 입장, 양자가 동등한 지위를 가진다는 입장, 그리고 단칭 인과는 어떠한 규칙성이나 법칙 과도 무관하게 성립한다는 입장이 있을 수 있다.

이러한 각각의 가능한 입장들에 대한 평가는 지금 이 대목에서는 중요하지 않다. 지금 우리에게 중요한 것은 적어도 어떤 단칭주의적 입장은 일반주의보다 더 나은 대안일 수 있다는 점이다. 위에서 보았듯이, 일반주의와 단칭주의를 대립시키는 가장 큰 쟁점은 과연 개별적인 인과 관계가 경험 가능한가의 여부였다. 이 쟁점과 관련하여 흄적인 생각이 설득력이 없다면 일반주의 역시 매력을 상실하게 된다.¹⁴⁾ 그렇다면 ‘인과의 법칙적 성격 원칙’도 굳이 받아들일 필요가 없게 될 것이며, 무리하게 개별자 동일성을 고수할 이유도 사라질 것이다.

1.4. 유형 부수현상론

개별자 동일성을 포기한다면 데이빗슨의 입장은 수반 논제만을 받아들이는 전형적인 비환원적 물리주의와 유사하게 된다. 특이하게도 데이빗슨 약수반 개념을 이용하여 자신의 이론을 설명하는데(Davidson 1993: 4, n4), 그가 왜 그렇게 하는지, 그렇게 하는 것이 적절한지에 대해서는 의문의 여지가 많다. 어쨌든 여기서는 다른 일반적인 물리주의자들의 견해를 따르도록 하겠다.

13) 관련된 논의들에서 어떤 철학자들은 ‘비환원적’이라는 말과 ‘원초적’이라는 말을 교환 가능한 것으로 사용한다. 반면에 ‘원초적’이라는 말에 비환원적이라는 것보다 더욱 강한 의미를 부여해서, 인과적 개념이 (인과적 개념을 포함한) 어떠한 개념으로도 환원 불가능하다고 사용하는 철학자들도 있다. 이 경우 원초주의는 비환원주의를 함축하지만, 그 역은 아니다(Carroll 2009: 281).

14) 양상 개념에 대한 최근의 태도를 보면 흄적인 생각이 퇴색하고 있다고 보아도 좋을 것 같다.

주로 총체적 수반 개념을 통해 정식화되는 물리주의에 따르자면, (미시) 물리적 사실들이 고정되면 모든 가능 세계에서 세계의 다른 모든 사실들 역시 필연적으로 고정된다. 혹은 가능한 모든 세계에서 물리적 차이 없이 다른 차이도 없다. 수반 논제는 다른 사실들의 환원 불가능성과 양립 가능하며, 정신적 개별자에 정확히 대응하는 물리적 개별자가 있어야 한다는 요구도 하지 않는다. 달리 말하자면, 수반 논제는 두 세계가 정신적 기술에서 차이가 있을 경우 물리적 기술에서도 차이가 있어야 한다고 말할 뿐, 어떤 물리적 차이에 의해 어떻게 정신적 차이가 생기는지 구체적으로 밝혀야 한다는 요구를 하지 않는다.

정신 인과 문제와 관련하여 수반 논제가 진정으로 말하고자 하는 바는 미시 물리적 속성이 인과에 있어서 기초적인 속성이라는 것, 다시 말해서 모든 인과력은 미시 물리적 존재자들로부터만 온다는 것이다. 그런 면에서 물리주의자들이 ‘수반’이라는 말에 부여하는 의미는 ‘형이상학적 결정’이나 ‘근거’라는 말로 더욱 잘 포착될 수 있다(Child 1994: 196-7). 그런데 이는 달리 말하자면, 단순히 두 종류의 속성들 사이의 공변 관계만을 논하는 수반 개념만 가지고는 물리주의를 정식화하기에 불충분하다는 뜻이기도 하다. 그럼에도 불구하고 수반 개념은 최소한의 물리주의를 정식화할 때 요구되는 것으로 흔히 이야기되고 있다. 이런 관행에 따라 이 글에서는 ‘수반 논제’라는 말로 단순한 공변 관계 이상의 것, 그 이상 필요한 것이 무엇이든 물리주의를 정식화하기에 충분한 것을 의미하고자 한다.

이제 데이빗슨을 비롯한 비환원적 유물론자들에게 공통적으로 가해지는 부수현상론 비판을 보자. 이 비판은 개별자 물리주의가 제대로 된 동일성 기준을 제시하지 못하고 있다는 비판과 독립적으로 전개되는 더욱 강력한 비판이다. 맥러플린(McLaughlin 1989)에 따르면, 부수현상론은 두 종류로 나눌 수 있다.

하나는 유형 부수현상론이다. 유형 부수현상론에 따르면, 사건들은 물리적 유형에 속하는 덕분에(in virtue of) 원인이 될 수는 있지만, 정신적 유형에 속하는 덕분에 원인이 될 수는 없다. 다른 하나는 개별자 부수 현상론이다. 개별자 부수현상론에 따르면, 물리적 사건은 정신적 사건을 야기

할 수 있지만, 정신적 사건은 아무것도 야기하지 못한다.

이들 중에서 비환원적 유물론에 가해지는 비판이 유형 부수현상론이라는 점은 분명하다. 데이빗슨의 개별자 동일성 이론을 떠올려 보자. 부수현상론 비판에 직면해서, 데이빗슨은 자신의 철학에서 정신적 사건이 물리적 사건의 원인이 될 수 있음을 몇 번이고 강조한다. 하지만 지난 절에서 보았듯이, 정신적 사건이 물리적 사건의 원인이 되는 것은 그 정신적 사건과 동일한 물리적 사건이 엄격한 법칙의 뒷받침을 받는 덕분이다. 그러면 정신적 사건은 물리적 기술이 주어지는 덕분에, 즉 물리적 속성 덕분에 인과력을 가진다는 말이 아닌가?¹⁵⁾

이와 같은 부수현상론 비판의 범위가 단지 정신적인 것의 인과력에만 한정되지 않는다는 점을 지적해 둘 필요가 있다. 어떤 이론 혹은 인과적 설명 유형이 물리학으로 환원되지 않는다고 주장하면서 동시에 수반 논제를 받아들인다면, 그 이론 혹은 인과적 설명 유형에서 거론하는 속성들은 모두 유형 부수현상론의 위협에 시달리게 된다. 이 문제는 이론 혹은 인과적 설명 유형의 개수가 많을수록 심각하게 다가온다. 그러면 우리는 실제로 얼마나 많은 이론 혹은 인과적 설명 유형을 가지고 있는가? 정확히는 모르겠지만 꽤 많을 것이라고 생각할 만한 이유들이 있다.

나는 우리가 가진 이론들 내지는 개별 학문들의 수와 정확히 같은 수의 근본적 속성들이 있다고 생각하지는 않는다. 어떤 사람들은 계층화된 속성들의 위계에 더해서 각각의 계층에 정확히 대응하는 학문들이 있다고 생각하기도 하는데, 이는 매력적인 그림이기는 하지만 유보 없이 그대로 받아들이기에는 지나치게 경직되어 있는 그림이기도 하다. 나는 차라리 참된 인과적 설명 유형들의 수만큼 많은 속성들이 있다고 말하는 게 낫다고 생각한다.

이처럼 이론 혹은 개별 학문과 인과적 설명 유형을 구분하고 후자를 선

15) 맥러플린(McLaughlin 1989)은 그렇다고 해서 엄격한 법칙만이 유일하게 인과적 속성을 부여한다고 해석할 필요는 없다고 함으로써 데이빗슨을 구원하고자 한다. 그에 따르면 엄격하지 않은 법칙도 인과적 속성을 줄 수 있다. 하지만 그렇다고 하더라도 역시 그 엄격하지 않은 법칙을 포섭하는 엄격한 법칙이 있어야 한다는 제약 조건은 남아 있다는 점에서 부수현상론 문제를 완전히 해결했다고 볼 수는 없으며, 그 자신이 그럴 의도를 가지고 있었던 것 같지도 않다.

호하는 몇 가지 이유가 있다. 우선 어떤 인과적 설명 유형들을 과연 이론이라고 할 수 있을지가 문제이다. 예를 들어, 믿음과 욕구 내용에 의해 행동을 설명하는 지향적 심리학에 대해 그것을 일종의 이론으로 보려는 시도들도 있지만, 그렇게 볼 수 없다고 생각하는 사람들도 있다. 후자에 속하는 사람들은 지향적 심리학이 그럼에도 불구하고 설명과 예측을 매우 잘 해내고 있다는 데에는 동의한다. 그렇다면 어떤 인과적 설명이 잘 작동하기 위해 반드시 이론의 일부일 필요는 없을 수도 있다는 결론이 나온다. 이런 생각까지 포괄하려면 이론이나 개별 학문보다는 중립적인 인과적 설명 유형에 대해 말하는 것이 나을 것이다.

또, 개별 학문들이 반드시 단일한 유형의 인과적 설명들로 이루어져 있다고 볼 이유가 별로 없다. 생물학을 예로 들어보자. 생물학에는 유전학, 분자 생물학, 형태학, 분류학, 생태학 등등의 다양한 분야들이 있다. 이러한 분야들이 모두 단일한 유형의 인과적 설명 유형으로 되어 있다고 보이지는 않는다. 다른 분과 학문들도 정도의 차이는 있겠지만, 모두 여러 유형의 인과적 설명들을 동원하고 있는 것으로 보인다. 현재 물리학조차도 상대성 이론을 주축으로 하는 거시적 대상들에 관한 인과적 설명과 양자역학을 주축으로 하는 미시적 대상들에 대한 인과적 설명으로 나뉘어 있다. 하나의 개별 학문은 오히려 관심사에 의해 여러 인과적 설명 유형들이 느슨하게 묶여 있는 형태로 되어 있다고 해야 할 것이다. 따라서 개별 학문들은 새로운 경험적 발견, 관심사의 변화 등등 여러 이유에 의해 분과 내적으로나 분과들 사이에서나 다양한 수준의 재편을 겪을 가능성이 얼마든지 있다. 그것들의 수가 인과적 설명 유형의 수와 일치하지 않는다고 하더라도 이상하지 않다.

그렇다면 인과적 설명 유형들은 무엇에 의해 나뉘는가? 미시 물리적 존재자로부터 시작해 분자, 세포, 기관, 개체, 그리고 사회 등으로 나아가는 식의 계층화된 세계상은 우리에게 주요 인과적 설명 유형들을 분류하는 한 가지 기준을 제공하는 것처럼 보인다. 이런 그림에 의한 인과적 설명 유형들의 분류는 개별 학문들에 정확히 대응하는 분류라기보다는 오히려 인간이 세계를 이해하는 한 가지 그럴듯한 방식에 따른 분류일 것이다.

물론 이런 분류 체계에서 각각의 층위에 해당하는 전형적이고 대표적인 분과 학문이 있다고 보는 것은 가능하다.

어쨌든 유형 부수현상론의 충격을 완화하려면 인과적 설명 유형의 수는 적을수록 좋을 것이다. 하지만 데이빗슨을 포함하는 비환원적 물리주의자들은 일반적으로 이론들 혹은 인과적 설명 유형들이 우리가 흔히 생각하는 것과 괴리될 정도로 많이 줄어들 것이라고 생각할 이유를 가지기가 어려운 사람들이다. 우선 그들은 모두 정신적 유형이 물리적 유형으로 환원되지 않는다는 데 동의한다. 여기에 더해서, 데이빗슨의 경우 간단히 언급만 하고 자세히 논의를 하지 않기 때문에 종종 간과되기 쉽지만, 생물학을 비롯한 몇몇 개별 과학들이 물리학으로 환원되지 않는다고 생각한다고 말한 바 있다(데이빗슨 1974/2012d: 393). 그가 이렇게 생각하는 이유는 쉽게 짐작할 수 있다. 생물학에서는 그 나름의 느슨한 법칙들에 의한 설명이 이루어지겠지만, 물리학에서와 같이 엄격한 법칙을 기대할 수는 없을 것이다. 그렇다면 데이빗슨에게는 적어도 심리학, 생물학, 물리학이라는 세 영역이 환원불가능하게 존재하는 것이다. 다른 비환원적 물리주의자들 역시 복수실현화가능성 등등의 체계적인 이유로 유형들 간 환원불가능성을 주장하는데, 이러한 체계적인 이유들이 단지 정신적인 것과 물리적인 것 사이에서만 작동한다고 보기는 어렵다.

더욱 중요한 점은 현재 개별 학문 종사자들에게 이론 간 환원, 특히 물리학으로의 환원이 그렇게까지 큰 의미를 가지지 못하는 것처럼 보인다는 점이다.¹⁶⁾ 비환원적 물리주의는 이와 같은 세태를 반영하는 입장이기 때문에 현재 주류인 물리주의 중에서도 주류적인 입장이 될 수 있었던 것이다. 따라서 데이빗슨을 포함하는 비환원적 물리주의자들은 일반화된 유형 부수현상론에 직면하게 된다. 유형 부수현상론이 말하고자 하는 것은 대량의 인과적 설명들이 거짓이라는 것이 아니라, 형이상학적 관점에서 보면 2등급 참이라는 것이다. 사건들은 오직 미시 물리적 유형에 속함으로써만

16) 뒤프레(Dupré 2010: 290, n2)에 따르면, 가장 단순한 화학적 물질인 수소 원자가 가지는 성질이 물리학으로 완전히 환원되느냐에 대해서만 논란이 있다. 그 이상의 복잡성을 가지는 대상에 대해서는 논란의 여지가 없다고 한다. 그는 환원은 실제 과학과 무관한 철학적 논제라고 보고 있다.

원인이 될 수 있다. 다른 유형들은 적어도 형이상학적 관점에서 보면 인과적으로 무력하다.

유형 부수현상론 비판에 대한 데이빗슨 자신의 답변은 만족스럽다고 보기 힘들다(Davidson 1993: 13). 그는 인과란 구체적 개별자인 사건들 사이의 외연적 관계이므로, 그것이 어떻게 기술되든지, 즉 어떤 속성을 가지는지와 상관 없이 성립한다고 주장한다. 하지만 A와 B사이의 어떤 관계가 외연적이라 하더라도 그것이 A와 B가 가지는 어떤 속성 때문에 성립한다고 볼 수 있는 경우가 대부분이다. 예를 들어, A가 B보다 무게가 가볍다는 외연적 관계는 A의 무게와 B의 무게라는 속성 때문에 성립한다고 볼 수 있다(McLaughlin 1993: 33-34).¹⁷⁾ 적어도 일상적 언어 관행에 비추어 보면 이런 비판은 매우 그럴듯하게 들린다. 물론 일상적 언어 관행은 오도적일 수도 있고, 다르게 해석될 여지도 있다. 그래도 어쨌든 데이빗슨에게는 우리가 가지고 있는 자연스러운 직관이 실제로는 철학적 중요성을 가지지 않는다는 점을 해명해야 할 과제가 있을 텐데, 그는 그렇게 하지 않고 있다.

데이빗슨의 옹호자들 중에는 데이빗슨이 속성에 관해 유명론자였다는 점에 주목하여 부수현상론 비판에서 벗어나려고 하는 논자들도 있다.¹⁸⁾ 여기서 유명론은 결국 속성이 그것에 대한 우리의 인식과 독립적인 실재성을 가지지 않는다는 뜻이므로, 반실재론이라 이해해도 괜찮을 것이다. 이러한 해석에 따르면 부수현상론 비판은 만약 정신적 속성이 있다면 그것은 인과적으로 무력하다는 비판이므로, 이 비판의 전건을 거부하는 데이빗슨의 철학에는 해당되지 않는 비판인 셈이 된다.

데이빗슨의 철학을 위와 같은 식으로 해석하면 몇 가지 의문이 생겨난다. 우선 속성은 그것이 부여되어 있는 대상을 개별화하고 이해하는 데에 결정적인 역할을 한다고 생각되는 것이 보통이다. 속성이 반실재적이라면 대상들은 어떻게 되는가? 두 가지 가능한 답변이 있을 것이다. 대상들은

17) 차일드(Child 1994: 189)는 기본적으로 데이빗슨의 답변을 따르고 있는데, 어떤 이유에서인지 시공간적 관계는 속성에 의해 설명되지 않는다고 여기는 듯하다.

18) Robb and Heil(2014: 5.5) 참조. 국내에서는 특히 백도형(2014)이 대표적으로 이러한 노선을 취하고 있다.

속성과 무관하게 개별화되거나, 아니면 속성과 마찬가지로 반실재적일 것이다. 둘 중에서는 전자가 더욱 안 좋은 선택지일 것이다. 우리는 속성을 통해서 대상을 파악한다고 보통 생각하는데, 대상과 그것을 언급하는 사건이 속성과 무관하게 개별화되어 있다면, 우리는 대상 그 자체에 대해 무엇을 통해 아는가? 이 경우 대상은 신비한 것이 된다. 반면에 후자를 택한다면 우리는 총체적인 반실재론을 얻게 되는데, 그러면 적어도 대상의 가지성은 보장을 받게 된다.¹⁹⁾ 대신 이 경우 인과 관계는 속성과 무관하게 성립한다기보다는 속성과 무관하게 성립하는 인과 관계라는 개념 자체가 사라진다고 보아야 한다.

어느 쪽을 택하든 다시 다음과 같은 의문이 제기될 수 있다. 속성 유명론은 도대체 어떤 의미에서 물리주의인가? 대상들은 속성과 무관하게 개별화되어 있으므로, 물리적이지도 정신적이지도 않거나, 아니면 대상 그 자체는 존재하지 않으므로 물리적 속성을 가진 존재자들이 더욱 근본적인 존재자들과 말할 이유가 없을 것처럼 보이기 때문에 묻는 것이다. 물론 이 대목에서 물리주의와 결별할 수도 있겠지만, 데이빗슨의 철학과 더욱 잘 조화되는 것처럼 보이는 다른 답도 주어진 바 있다(백도형 2014: 391-95). 그에 따르면 속성 유명론은 물리학이 세계에 대한 가장 포괄적인 설명을 제공한다는 의미에서 물리주의적이다. 예컨대, 모든 정신적 대상 혹은 사건에 대응하는 물리적 설명은 있지만, 그 역은 아니라는 것이다.

이런 물리주의는 기존에 우리가 알던 물리주의에 비해 훨씬 온건한 입장인 것처럼 보이지만, 그럼에도 불구하고 여전히 문제적인 측면이 있다. 앞서 우리는 정신적 사건과 물리적 사건 사이의 동일성 기준을 제시하는데 데이빗슨이 실패했음을 살펴보았다. 이는 임의의 정신적 사건에 대응하는 물리적 사건이 무엇인지 말할 수 없다는 뜻이기도 하다. 그런데도 물리학이 가장 포괄적인 설명을 제공한다고 할 수 있는가? 이 물음에 대해 아마도 다음과 같은 답이 주어질 것이다. 어쨌든 정신적 사건도 시공간상에서 일어나는 사건이라는 점은 분명하다. 정신적 사건의 개별화가 어떤 식

19) 그래서 백도형(2014: 379)은 후자를 선택하고 있다.

으로 이루어지든 물리학은 모든 경우에 그것에 대응하는 설명을 원리적으로 제공할 수 있을 것이라는 식으로 말이다. 하지만 우리가 현재 물리학의 이론과 실천을 주의깊게 관찰하면, 물리학이 원리적으로라도 가장 포괄적인 설명을 제공할 수 있으리라는 생각은 근거가 없다는 점이 드러날 것이다. 이는 3장에서 본격적으로 다루어질 내용이지만 미리 결론만 말하자면, 물리 이론의 조건부적 성격을 제거할 수 없기 때문에 그러하다. 물리학이 가장 포괄적인 설명을 제공한다고 하는 말이 그럴듯하게 들린다면, 그것은 모든 것이 모종의 물질로 이루어져 있다는 생각이 자연스럽게 받아들여지기 때문일 것이다. 하지만 그러한 생각과 물리학이 가장 포괄적인 설명을 제공한다는 생각은 별개의 것이다. 이 둘 사이에서 혼동하지 말아야 한다.

1.5. 결론

이렇게 데이빗슨의 무법칙적 일원론은 정신적, 물리적 개별자들 사이의 동일성 기준을 제시하지 못한다는 한 가지 문제와 유형 부수현상론에 빠지게 된다는 다른 한 가지 문제를 노출시키고 있다. 전자는 ‘인과의 법칙적 성격 원칙’을 채택하지 않고, 개별자 물리주의를 포기하면 쉽게 해결할 수 있지만, 후자에 대해서는 데이빗슨의 철학 내에서 해법을 발견하기가 어렵다. 유형 부수현상론은 비환원적 물리주의 전반에 적용되는 심각한 문제이므로, 더욱 자세히 검토될 필요가 있다.

2. 비환원적 물리주의와 배제 문제

이전 장에서 우리는 부수현상론이나 평행주의 같은 불편한 입장으로 귀결하는 것을 피하고자 한다면, 정신적인 것이 정신적인 유형으로서 인과력을 가진다는 것을 주장해야 한다는 점을 확인했다. 비환원적 물리주의를 확립시키는 데 크게 기여한 데이빗슨이 결국 유형 부수현상론 혐의로부터 자유롭지 못하다는 것을 우리는 보았는데, 이번 장에서는 그 이후의 여러 비환원적 물리주의자들 역시 그러하다는 점을 드러내려고 한다.

비환원론 물리주의자들은 수반 논제를 통해 정식화되는 물리주의와 정신적 속성의 인과력을 동시에 유지할 수 있다고 주장한다. 그런데 보다 강경한 물리주의자들은 비환원적 물리주의를 포함하는 그 이상의 여러 입장들을 공격하기 위한 강력한 논증을 가지고 있다. 그 논증이 제기하는 문제는 ‘배제 문제’라고 불리는데, 이 문제는 부수현상론 문제를 포괄하는 좀 더 일반적인 문제이며, 사실상 정신 인과 문제의 핵심이다. 우선 나는 배제 문제가 정확히 어떤 것인지 살펴본 뒤, 데이빗슨 이후에 제안된 주목할 만한 비환원적 물리주의들이 이에 대해 만족스러운 답을 주지 못했다는 점을 보일 것이다. 여기에 더해서 비환원적 물리주의는 배제 문제를 해결하려는 과정에서 독립적으로 설득력이 있는 외재주의와도 조화되기 어렵다는 관련된 약점 또한 드러내 보인다.

2.1. 배제 문제

배제 문제는 흔히 다음과 같이 정식화된다.²⁰⁾

구별됨: 정신적 속성 혹은 사건은 물리적 속성 혹은 사건과 구별된다.

20) 여기서 나는 기본적으로 베넷(Bennett 2008)의 정식화를 따르고 있다.

효력: 어떤 정신적 사건은 물리적 사건을 일으키고, 그것의 정신적 속성 덕분에 그렇게 한다.

폐쇄성(완결성): (원인이 있는) 모든 물리적 사건들은 충분한 물리적 원인을 갖는다.

비과잉결정: 체계적인 인과적 과잉결정은 없다.

배제: 어떤 결과도 그것이 과잉결정되어 있지 않는 한, 하나보다 더 많은 충분한 원인을 가지지 않는다.

이러한 주장들을 모두 동시에 수용하는 것은 비일관적이다. 정신적 속성이 설령 그 자체로서 인과력을 가진다고 하더라도, 물리적 사건들이 완전히 닫혀 있는 영역을 이루고 있기 때문에 인과적 역할을 하고자 한다면 과잉결정을 일으킬 수밖에 없다. 달리 말하자면, 정신적 속성은 별도의 인과력을 가지고 있다고 하더라도 정당하게 할 수 있는 일이 없게 된다. 그 점에서 이 논증은 정신적 속성이 별도의 인과력을 가지지 않는다는 점을 보이게 하는 부류의 논증들과는 접근 방식이 약간 다르다.

논의를 진행하기 전에 몇 가지 짚고 넘어가야 할 부분이 있다. 양자 역학의 발전 덕분에 우리는 물리적인 원인과 결과 사이의 필연성을 더 이상 자명한 것으로 받아들이지 않는 시대에 살고 있다. 완결성에서 ‘충분한’이 ‘결과를 필연적으로 낳는’을 의미하는 것이라면, 곧바로 물리 영역의 인과적 폐쇄성이 성립하지 않는다는 것으로 보고 배제 문제를 일소할 수도 있지 않을까? 그렇지 않다. 비결정론을 받아들인다면, 충분한 원인이라는 것은 결과를 필연적으로 낳지는 않지만, 어쨌든 결과를 일으키기에는 충분한이라는 의미를 가지게 될 것이다. 즉, 물리적 원인은 가능한 물리적 결과의 선택지들을 제공할 것이며, 물리적 결과로부터 원인으로 거슬러 올라가면 우리는 어쨌든 물리적 원인들만을 마주치게 될 것이다. 따라서 논의의 편의를 위해 마치 우리가 결정론적 우주에 살고 있는 것처럼 생각하더라도 문제는 없다.²¹⁾

21) 로위(Lowe 2006: 19)는 비결정론에만 지나치게 의존하여 인과적 폐쇄성을 거부하려는 것처럼 보인다.

과학 이론의 발전과 관련하여 더욱 심각하게 생각해 볼 문제는 폐쇄성이 어떤 수준에서 성립하느냐는 것이다. 관련된 논의들에서 폐쇄성은 최종적인 미시물리학의 영역에서 성립하는 것처럼 흔히 말해지고 있다. 그런데 만약 최종적인 이론이 없다면 어떻게 되는가?

무한한 하강은 모든 것이 거시적이라는 바로 그 이유 때문에 거시적 세계에서 편안한 평등한 존재론적 태도를 낳는다. 중간자들, 분자들, 마음들, 그리고 산들은 모든 의미에서 존재론적으로 동등하다. 왜냐하면 인과력을 위한 특권화된 장소가 있을 수 없기 때문에, 그리고 그것들은 어딘가에 있어야만 하기 때문이다. 그것들은 모든 곳에 있다. 그래서 무한한 하강은 자연 전체에 명예와 권력을 주는 평등한 형이상학을 낳는다(Schaffer 2003: 512-13).

이처럼 최종적인 수준은 없으며, 모든 수준이 거시적이라는 것은 쉐퍼 자신이 선호하는 대안이다. 하지만 그는 다른 가능성 또한 제시하고 있다. 최종 이론도 없고 무한한 하강만이 있지만, 중간 어딘가에 수반 기저가 있을 수도 있다. 그 수준을 기준으로 그 위, 그리고 그 아래의 다른 모든 수준들은 그 수준에 수반한다. 그러면 그 위의 수준에서는 상향식 인과만이, 그 아래의 수준에서는 하향식 인과만이 있을 것이다. 폐쇄성 역시 그 수반 기저 수준에서만 성립하는 것이라고 볼 수 있다. 물리주의자들의 입장에서 쉐퍼 식의 평등한 그림이든, 이러한 식의 어중간한 그림이든 달가워하지 않겠지만, 그래도 후자를 택할 수밖에 없을 것이다.

비과잉결정이라는 전제에서 “체계적인” 과잉결정이 없다고 말하는 이유는 특수한 경우의 과잉결정은 받아들일 준비가 되어 있기 때문이다. 두 명의 사수가 발사한 두 발의 총알이 동시에 한 사람에게 적중한 경우, 각각의 총알은 사람을 죽이기에 독립적으로 충분한 원인이 되기 때문에 우리는 그 결과가 과잉결정되어 있다고 말한다. 체계적인 과잉결정이 없다는 것은 정신적 속성이 언제나 이러한 식으로 물리적 결과를 과잉결정하는 것은 이상하다는 것이다.

어쨌든 비일관성을 피하고자 한다면, 위의 주장들 중 하나는 거부해야

한다는 점은 분명하다. 이제 비환원적 물리주의자들이 어떤 전략을 취하는지 살펴해보도록 하자.

2.2. 체계적인 과잉결정의 가능성 혹은 불가능성

최근의 주목할 만한 비환원적 물리주의자들 가운데 한 사람인 베넷(Bennett 2003, 2008)은 배제, 즉 어떤 결과도 그것이 과잉결정되어 있지 않는 한, 하나보다 더 많은 충분한 원인을 가지지 않는다는 주장을 거부하는 해결책을 제시한다. 그렇다면 어떤 물리적 결과에 하나보다 더 많은 충분한 원인이 있지만, 그것이 과잉결정을 일으키지는 않는다는 것이다. 이것은 비과잉결정성을 거부하는 것으로 해석될 수도 있다. 그 경우 정신적 속성은 물리적 결과를 체계적이지만 문제 없는 방식으로 과잉결정하는 셈이다.

우리가 흔히 떠올리는 과잉결정의 사례를 생각해 보자. 여기서 두 명의 사수가 발사한 각각의 총알들은 ‘독립적으로’ 충분한 원인이라는 점이 중요하다. 베넷이 보기에 정신적 원인과 물리적 원인은 이렇게 각각 독립적으로 충분한 원인들이 아니다. 그 둘은 긴밀한 관련을 맺고 있다. 베넷은 정신적 원인과 물리적 원인이 정말로 과잉결정을 일으킨다면, 다음의 두 반사실적 조건문이 둘 다 공허하지 않게 참이어야 한다고 주장한다.

(O1) p 없이 m 이 일어났더라도, e 는 여전히 일어났을 것이다.

(O2) m 없이 p 가 일어났더라도, e 는 여전히 일어났을 것이다.

(Bennett 2008: 288. 여기서 m 은 정신적 원인, p 는 물리적 원인, e 는 결과임)

(이 반사실적 조건문들은 인과 관계에 대한 분석으로 제시된 것이 아니라 그저 추론을 돕기 위해 제시된 것이다. 인과 관계의 본성에 대해 베넷

은 중립적인 태도를 취하고 있다.)

베넷은 (O1)은 정신적 속성의 효력을 거부하지 않는 한, 공허하거나 거짓일 수 없을 것이라고 본다. 즉, (O1)은 공허하지 않게 참이다. (그런데 베넷은 물리주의자로서 물리적 원인 없이 정신적 원인이 있다는 말을 어떤 식으로 이해하고 있는가?) 그녀가 문제삼는 것은 (O2)이다. 비환원적 물리주의자가 보기에 (O2)는 공허하거나 거짓이다.

(O2)가 공허하지 않은 경우에 대해 생각해 보자. m 과 p 가 각각 e 를 일으키기에 충분한 원인이지만, 수반 논제에 따라 p 가 m 을 필연화한다면, m 없이 p 가 일어난다는 것은 불가능하므로 (O2)는 공허하게 될 것이다. 따라서 p 가 m 을 필연화하지 않아야 (O2)는 공허하지 않게 된다. 놀랍게도 베넷은 p 가 m 을 필연화하지 않는 경우가 흔하다고 주장한다. 관련된 논의에서 우리가 흔히 p 에 해당하는 것으로 거론하는 뉴런 활동의 패턴이라든지, C-섬유의 발화 같은 것들은 특정한 배경 조건이 주어질 때에만 정신적 속성을 필연화하기 때문이라는 것이다. 그러면 (O2)의 사례인 “고통 없이 C-섬유가 발화했더라도, 나는 여전히 난로로부터 황급히 손을 뺏을 것이다”와 같은 명제는 거짓이 되는데, 왜냐하면 고통 없이 C-섬유가 발화하는 경우와 고통과 함께 C-섬유가 발화하는 경우는 서로 다른 배경 조건 하에 있기 때문이다. 내가 난로로부터 황급히 손을 떼는 결과가 일어나는 상황, 즉 고통과 함께 C-섬유가 발화하는 상황과 다른 배경 조건 하에서 C-섬유가 발화했다면, 그것은 확실히 동일한 결과를 보장할 수 없을 것이다. 따라서 이런 경우 (O2)는 거짓이라고 베넷은 주장한다.

그런데 C-섬유의 발화가 내가 난로로부터 황급히 손을 떼는 결과를 보장하지 못한다면, 그리고 충분한 원인이라는 것이 결과를 필연적으로 일으키는 것으로서 전제되어 있다면, 내가 손을 떼는 결과를 보장하지 못하는 C-섬유의 발화는 애초에 충분한 원인이 아니지 않을까? 이 물음에 대해 베넷은 다음과 같이 답한다.

기본적으로, 만약 당신이 인과적 충분성을 엄격한 충분성의 일종으로 생각한다면, 오직 일상적 사건들, 배경 조건들, 인과적 중간 단계들과 같은 것들의 크고 복잡한 총합만이 무언가를 위해 인과적으로 충분하다고 셈해지는

데, 그러면 행위를 위해 도대체 인과적으로 충분한 유일한 물리적 사건들은 복잡한 비국지적 사건들일 것이다. 이것들은 정신적인 것들을 보장하고, 따라서 (O2)의 모든 사례들은 공허한 것으로 드러난다. 만약, 반면에, 당신이 인과적 충분성을 그저 상황들 속에서의 충분성으로 생각한다면, 보다 일상적이고, 국지화된 물리적 사건들과 속성들이 행위를 위한 인과적으로 충분한 것으로 섰해질 것이고, (O2)의 모든 사례들이 공허하지는 않을 것이다.

따라서 심지어 물리주의자조차도 (O2)의 공허하지 않은 사례들이 있다고 완벽하게 잘 말할 수 있다(Bennett 2008: 290-91).

결국 베넷의 요점은 결과를 보장하지 못하는 원인도 어떤 의미에서는 충분한 원인이라는 것이다. 하지만 나는 어떤 의미에서 결과를 보장하지 못하는 원인도 ‘충분한 원인’이라고 불릴 수 있다는 점이 지금 우리의 논의와 정말로 관련이 있는 것인지 의심스럽다.

원인을 결과가 일어나기 위한 필요충분조건으로 분석하는 입장에서, 실제로 어떤 결과의 완전한 원인을 명시적으로 제공하려는 시도는 현실적인 어려움에 부딪힐 수밖에 없다. 무엇보다도 원인과 배경 조건 사이의 선명한 구분을 제시하는 것이 쉽지 않기 때문에 그러하다. 그렇기 때문에 물리주의자들도 어떤 맥락에서는 필요에 따라 ‘충분한 원인’이라는 말을 느슨한 의미로 사용할 수 있을 것이다. 하지만 배제 문제에 관한 논의에서 ‘충분한 원인’이 그렇게 느슨한 의미로 사용되는 경우는 없는 것 같다. 물리 영역의 인과적 폐쇄성에 대한 정식화에서 ‘충분한 원인’이 엄격한 의미에서의 충분한 원인이 아니라면, 그것은 비환원적 물리주의자들을 포함하는 그토록 많은 철학자들을 괴롭히지 못했을 것이다. 그것이 주장하는 바가 그저 물리적 사건이 특정 상황 속에서의 충분한 원인을 가진다고 하는 것이었다면, 그리고 그것이 우리가 주장할 수 있는 전부라면, 엄격한 의미에서는 물리 영역의 인과적 폐쇄성이 성립한다는 보장이 없으므로, 문제적인 과잉결정이 정말로 발생한다는 것을 강고한 물리주의자들이 다른 입장을 가진 논자들에게 납득시키기가 힘들어질 것이다. 그러면 지금처럼 배제 문제가 심각하게 다루어지지 않았을 것이다. 따라서 여기서 ‘충분한 원인’은 경우에 따라 느슨한 의미로 쓰이는 것처럼 보일지라도 엄격한 의미의

생략형으로 이해되는 것이 옳다.

그렇다면 이 문제와 관련해서는, 물리주의자의 입장에서 (O2)는 결코 거짓일 수가 없으며, 오로지 공허할 수밖에 없다. 그래도 어쨌든 (O2)는 공허하므로 과잉결정은 발생하지 않는 것이고, 따라서 배제 문제는 해결되었다고 베넷은 주장할 것이다. 정말로 문제가 해결되었는가?

앞서 우리는 베넷이 배제 문제를 일으키는 다섯 개의 비일관적 주장들 가운데서 배제, 즉 어떤 결과도 그것이 과잉결정되어 있지 않는 한, 하나보다 더 많은 충분한 원인을 가지지 않는다는 주장을 거부하려 하고 있다는 점을 보았다. 그것은 어떤 결과가 하나보다 더 많은 충분한 원인을 가지고 있더라도, 과잉결정이 일어나지 않는다는 것을 뜻한다. 내가 주목하는 곳은 “하나보다 더 많은”이라는 구절이다. 베넷의 논증에서 물리적 사건은 정말로 하나보다 더 많은 충분한 원인을 가지는가? 정신적 원인은 물리적 원인과 정말로 구별되는가? 이 물음은 결국 베넷이 정신적 속성의 효력을 제대로 유지하고 있는지를 묻고 있는 것이다. 그녀의 그림에서 정신적 속성은 정말로 정신적 속성으로서 인과적 역할을 하는가? 아니라면 다시 유형 부수현상론으로 전락할 위험에 처하게 된다.

결국 베넷은 정신적 속성의 인과적 효력에 대해 말한 것이 거의 없다. 그녀도 이러한 비판을 예상은 하고 있다(Bennett 2008: 301ff). 그런데 이에 대해 그녀는 “배제 문제를 해결하는 것은 정신적 사건과 속성의 인과적 효력을 옹호하는 것을 요구하지 않는다”(Bennett 2008: 281)는 이해할 수 없는 답변만을 내놓고 있다. 만약 정말로 그렇게 생각한다면 어째서 배제 문제를 위와 같이 정식화하고 복잡한 답변을 내놓았는가? 진심으로 정신적 속성의 인과적 효력을 옹호할 필요가 없다면 왜 굳이 복잡한 과정들을 통해 배제 문제를 해결하려 하는지가 의문이다.

베넷의 논의는 거의 전적으로 정신적 원인이 물리적 원인과 얼마나 밀접하게 연결되어 있는지, 정신적 원인이 물리적 원인에 의존적인지를 설명하는 데 할애되어 있다. 둘 사이의 밀접함, 의존성을 강조하다보면 반대로 정신적 원인의 독립성이 위협받고, 그러면 정신적 속성의 효력을 유지하기가 어려워지리라는 것이 자연스러운 생각의 발로이다. 베넷은 물리주의자

로서는 정신적 속성의 독립적 인과력에 대해 말할 수 없다는 점을 마침내 시인한다(Bennett 2008: 301). 하지만 그렇다면 정신적 속성이 어떻게 물리적 속성과 구별될 수 있는지가 의문이다. 우리는 일반적으로 속성들의 개별화가 그것들이 하는 인과적 역할에 따라 이루어진다고 생각하기 때문이다.²²⁾

베넷을 비롯한 비환원적 물리주의자들은 당연히 자신들의 그림 속에서도 정신적 속성의 효력은 유지된다고 주장해야 한다. 그리고 실제로도 비환원적 물리주의는 그런 주장을 하는 것으로 취급되어 왔다. 이들은 정신적 사건, 속성이 물리적 속성, 사건과 유형 동일성은 물론 개별자 동일성 관계에도 있지 않으며, 서로 구별된다고 말하고 싶어한다.²³⁾ 하지만 배제 문제에 직면하자 이들은 물리적 속성에 대한 정신적 속성의 ‘의존 관계’ 같은 새로운 구분을 도입하여 문제를 회피하려 한다. 이러한 새로운 구분을 부르는 명칭은 다양할 수 있다. 비환원적 물리주의자들은 ‘물리적 실현’에 대해 말하기도 하고(Shomaker 2013), 때로는 인과적 효력과 구분되는 설명력에 대해 말하는 ‘인과적 관련성’ 같은 개념을 도입하기도 한다(Audi 1993, Child 1994). 이러한 새로운 개념들이 공통적으로 하는 역할이 있다면, 그것은 속성들의 개별화와 인과적 효력 사이의 자연스러운 결합을 분리시키는 것이다.

비판적인 입장에서 보면, 위와 같은 식의 해결책은 미봉책에 불과한 것처럼 느껴진다. 나에게서는 베넷을 비롯한 비환원적 물리주의자들이 정신적 속성의 효력을 거부한 대가로 비과잉결정을 얻은 것처럼 보인다. 그러면 비환원적 물리주의는 속성들 간의 모종의 동일성을 주장하는 보다 강한 물리주의와 사실상 다를 바가 없게 된다. 베넷은 자신의 입장이 환원적 물리주의와 본격적인 이원론 사이에서 환원적 물리주의에 보다 가까운 중도적 입장이라고 말하는데(Bennett 2008: 286), 그런 입장이 유지되려면 정신적 속성이 물리적 속성과 인과적 효력이 아닌 무엇에 의해 구별될 수

22) 뎀포드(Mumford 2011, 2013)에 따르자면, 속성은 곧 인과력이다.

23) (Bennett 2003: 480)에서는 정신적 원인과 물리적 원인이 서로 동일하지는 않지만, 정확히 구별되지도 않는다고 말하기도 한다. 이는 더욱 이해하기가 어려운 말이다.

있는지를 보여주어야 할 것이다. 나는 그 가능성에 대해 회의적이다.

2.3. 내용 외재주의와 배제 문제

드레츠키(Dretske 1989, 1993)는 약간 다른 방향에서 정신 인과 문제에 접근하고 있다. 그가 가지고 있는 문제 의식은 심적 내용에 관한 외재주의가 옳다고 할 때, 그것이 정신 인과와 관련하여 일으키는 문제를 어떻게 해결할 것인가라는 물음에 맞춰져 있다. 또한, 그는 심적 내용이 내용으로서 행동에 인과적 역할을 한다는 것을 보여줄 수 있어야 한다는 생각을 가지고 있다(Dretske 1989: 3).

우선 심적 내용에 관한 외재주의를 간략히 설명하자면, 그것은 우리가 표상하는 내용이 개인 바깥에 있는 자연적, 문화적 환경에 적어도 부분적으로라도 의존적이라고 주장하는 입장이다. 흔한 예로, 내가 ‘물’을 생각하거나 의미하고 있을 때, 그것의 내용은 나의 내적 심리 상태만으로는 결정되지 않으며, H₂O라는 외적 현실과의 적절한 관계까지 고려해야 결정된다고 외재주의는 주장한다. 또, 내가 물이 H₂O라는 것을 직접 알고 있지는 못하므로, 물이라는 자연종에 관한 일차적인 지식을 가지고 있는 과학자들, 그리고 과학 지식이 전달되는 사회적 장치들과의 관계 역시 나의 심적 내용을 결정하는 데에 고려되어야 할 것이다.

그러면 외재주의는 어떻게 정신 인과와 관련하여 문제를 일으키는가? 문제는 우리가 행위에 관해 생각할 때, 행위가 순전히 개인 내적인 원인들에 의해 완전히 결정되어 나온다고 생각하는 경향이 있기 때문에 발생한다. 나의 내적 상태가 나의 행위를 완전히 결정한다면, 외적 요인은 불필요하게 될 것이다. 다음과 같은 비유를 생각해 보자.²⁴⁾ 자판기에 투입되는 동전이 가지고 있는 속성들 가운데서 동전의 화폐로서의 가치라는 속성은 조폐공사를 비롯한 사회적 관계에 의해 결정된다. 내가 그 동전과 정확히 같은 재질과 모양의 금속 덩어리를 만들어낸다면, 그것은 화폐로서의 진정

24) 이 비유는 Robb and Heil(2014: 7.1)에서 가져온 것이다.

한 가치를 가지지 않을 것이다. 하지만 그 금속 덩어리를 자판기에 넣으면, 자판기는 진짜 동전을 넣었을 때와 마찬가지로 음료수를 내놓을 것이다. 자판기의 행동을 결정하는 것은 투입되는 금속의 내적 속성들뿐이기 때문이다. 이러한 식으로 표상적 내용의 외적 속성은 행위와 인과적으로 무관한 것처럼 생각되기가 쉽다.

드레츠키는 “촉발하는(triggering) 원인”과 “조직하는(structuring) 원인”을 구분함으로써 이 문제를 해결하고자 한다. 그는 다음과 같은 직관적인 구분법을 제시한다. 내가 키보드에 있는 어떤 키를 누르면 화면에 있는 커서가 움직인다. 이때 내가 키를 누르는 행위가 화면에 있는 커서의 움직임을 촉발하는 원인이다. 그런데 내가 키를 눌렀을 때 커서가 움직이려면, 컴퓨터의 전원이 켜져 있어야 하고, 컴퓨터가 제대로 작동해야 하고, 프로그램이 제대로 되어 있어야 하는 등등의 여러 가지 배경 조건들이 필요하다. 이런 조건들이 갖추어져 있지 않다면, 내가 키를 눌러도 커서가 움직이지 않을 것이다. 그때 내가 찾는 이유가 커서의 움직임을 조직하는 원인이다. 촉발하는 원인은 조건이 갖추어져 있다면 결과를 보장하는 반면, 조직하는 원인은 촉발하는 원인의 유무에 따라 결과를 일으킬 수도, 일으키지 않을 수도 있다.

이상의 설명에서 보면, 조직하는 원인은 원인이 아니라 배경 조건과 동일시될 수 있을 것 같은데, 드레츠키는 원인과 배경 조건을 엄격하게 구분하지 않으며, 경우에 따라 조직하는 원인을 결과가 되는 사건을 일으키는 원인이라고도, 사건을 일으키는 데 필요한 배경 조건을 일으키는 원인이라고도 부른다. 또, 한 경우에서의 촉발하는 원인이 다른 경우에는 조직하는 원인이 될 수도 있다고도 말한다.

이제 어떤 표상적 상태 R^+ 가 있다고 해 보자. R^+ 은 내적 속성과 외적 속성을 가지고 있다. 내적 속성은 각종 신경생리학적 속성들일 것이고, 외적 속성은 특정 대상의 현전에 대한 신뢰할만한 내적 지시자이다. 외적 속성을 좀 더 쉽게 이해하려면, 온도계의 눈금과 기온 사이의 관계를 생각해 보면 된다. 온도계의 눈금은 그것에 대응하는 특정한 외부 온도를 가리키도록 되어 있다. 온도계의 눈금은 온도계의 내적 상태이면서, 외부 기온에

대한 정보를 반영하는 신뢰할만한 지시자로 작용하고 있다. (여기서의 지시, 혹은 반영 관계는 어쨌든 물리주의적으로 이해되기에 문제가 없는 관계라고 암묵적으로 가정되어 있다.) 또, 온도계는 우리의 표상이 그러하듯이 오작동을 일으킬 수도 있다. 물론 인공물이라는 점에서 우리의 표상 체계와는 중대한 차이가 있지만, 드레츠키는 그래도 이러한 유비를 통해 우리의 표상 체계도 이해할 수 있다고 생각한다.

R+가 등록하는 특정 대상이 호랑이, 혹은 호랑이처럼 보이는 대상이라고 해 보자. 다시 말해서, R+는 호랑이에 대한 내적 지시자이다. 호랑이의 출현은 R+의 개별자를 인과적으로 발생시키며, 이것은 내가 호랑이를 지각했다는 뜻이다. 그러면 R+라는 나의 내적 상태의 발생은 내가 몸을 피하는 행동을 일으키는 촉발하는 원인이 된다. 여기서 드레츠키는 R+이 나의 행동의 원인이 되는 것이 학습의 결과라고 말한다(Dretske 1993: 132). 학습이 일어나기 전이라면, 호랑이의 지각이 R+을 발생시키더라도, R+가 몸을 피하는 행동을 일으키지 않는다는 뜻이다. 또, 학습이 이루어진 다음에는 호랑이를 지각하지 않더라도 어떤 이유에서든(예를 들자면, 쌍둥이 지구에서 호랑이 비슷하게 생겼지만 포식자가 아닌 동물을 마주친 경우 등등) R+가 발생한다면, 몸을 피하는 행동이 일어난다는 뜻이기도 하다.

특정한 내적 상태의 발생이 촉발하는 원인이므로, 우리는 자연스럽게 외적 조건에 의존적인 표상적 내용이 조직하는 원인에 해당하리라고 예상할 수 있다. 드레츠키는 이 두 종류의 원인을 행동에 대한 서로 다른 종류의 설명과도 연결시킨다. 촉발하는 원인에 의한 설명은 ‘왜 내가 몸을 피하는 행동을 하였는가?’라는 물음에 답한다. 답은 ‘호랑이를 보았으므로’이다. 조직하는 원인에 의한 설명은 ‘왜 내가 호랑이를 보고 몸을 피하는 행동을 하였는가?’라는 물음, 즉 이유에 대한 물음에 답한다. 그 물음에 대한 답은 호랑이가 나의 포식자이기 때문일 수도 있지만, 의외로 내가 예전에 그 호랑이에게 빚을 진 적이 있는데 갚지 못해 면목이 없기 때문일 수도 있다. 이와 같은 식으로 외적 조건에 의존하는 심적 내용이 행동에 인과적으로 관련되어 있다고 드레츠키는 주장한다.

드레츠키의 이론은 종종 이중 피설명항(dual explanandum) 전략 가운데 하나로 분류되는데,²⁵⁾ 방금 논의한 내용에 비추어 생각해 보면 그렇게 분류하기가 곤란하다는 점을 알 수 있다. 이중 피설명항 전략은 간단히 말해서, 통상적으로 제시되는 배제 문제에서 각각의 충분한 정신적 원인과 물리적 원인이 하나의 결과를 두고 경합하는 그림이 애초에 잘못되어 있으며, 정신적 원인은 정신적 결과라는 피설명항을, 물리적 원인은 물리적 결과라는 피설명항을 각각 인과적으로 설명한다고 주장함으로써 문제를 해결하려는 전략이다. ‘포식자를 피함’이라는 행동에 대해 생각해 보자. 포식자를 보고 피하는 행동은 여러 가지 방식으로 이루어질 수 있다. 일반적으로 말해서, 하나의 행동은 여러 가지 신체적 움직임을 통해 구현될 수 있다. 이중 피설명항 전략에 따르면, 물리적 원인은 신체적 움직임을 인과적으로 설명하고, 정신적 원인은 행동을 인과적으로 설명한다. 물리주의적 틀 안에서 이와 같은 전략은 배제 문제와 비슷한 문제를 일으킨다. 정신적 원인은 행동을 일으키는 것으로 되어 있지만, 물리적 원인은 행동의 기저를 이루는 신체적 움직임을 일으키고, 신체적 움직임은 수반 논제에 의해 행동 또한 일으키게 되기 때문에 일종의 과잉결정이 발생하게 되는 셈이다. 그런데 위에서 보았듯이, 드레츠키는 심적 내용을 조직하는 원인으로 간주하고, 조직하는 원인은 결과를 일으키기에 충분한 원인이 아니기 때문에, 그가 전형적인 이중 피설명항 전략을 구사하는 것으로 보기는 힘들다. 비록 그가 행동과 신체적 움직임을 설명적 수준에서 구분하고 있더라도 말이다.

어쨌든 심적 내용의 역할에 대한 드레츠키식의 식의 설명은 몇 가지 의문을 불러일으킨다. 그는 학습이 구체적으로 어떻게 이루어지는지 거의 말해주지 않고 있다. 그와 같은 물리주의자의 입장에서 학습은 내적 상태가 변화되는 과정과 동일시될 것이다. 이 과정에서 외적 조건에 의존적인 심적 내용이 제대로 된 인과적 역할을 하지 않는다면, 결국 행동을 설명하는 데서 심적 내용은 배제될 것이다.

학습에 대한 설명이 잘 주어질 수 있다고 하더라도 정말로 심각한 문제

25) 베넷(Bennett 2007)도 이렇게 분류하고 있다.

는 여전히 남아 있다. 드레츠키의 설명에서 내적인 표상적 상태는 과거에 이미 완료된 학습의 결과로 행동을 일으킨다. 하지만 우리가 심적 내용이 행동의 원인이 되는 경우에 대해 자연스럽게 가지고 있는 생각은, 그것이 현재 우리의 행동에 인과적 효력을 행사한다는 것이다. 이런 생각에 비추어 보면, 드레츠키가 제시하는 설명은 심적 내용이 인과력을 행사하는 방식을 적절하게 그리고 있지 못하다는 인상을 받게 된다. 그것은 심적 내용이 행동의 직접적인 원인이 되지 못한다는 점에서 유형 부수현상론뿐만 아니라 개별자 부수현상론의 위험마저 안고 있다.

이는 행동에 대한 인과적 설명에서 촉발하는 원인과 조직하는 원인 사이의 구분이 드레츠키 자신이 앞서 제시했던 커서의 움직임에 대한 설명의 경우와 딱 들어맞지 않는 것처럼 보인다는 사실과도 관련이 있다. 커서의 움직임에 대한 설명에서 우리는 조직하는 원인이 문자 그대로 컴퓨터의 고정배선(hard-wiring)이라는 사실을 쉽게 파악할 수 있다. 그래서 심적 내용이 행동을 설명하는 경우에도, 내적인 표상적 상태의 발생이 촉발하는 원인이라면, 그 시점에서의 우리의 내적 상태, 학습에 의해 설치된 내적 배선이 조직하는 원인이 되어야 할 것 같다는 생각을 하게 된다. 그러면 결국 외재적 내용은 내용으로서는 현재 시점에서 이루어지는 행동에 대한 인과적 설명에서 배제되는 것이다. 그러면 애초에 배제 문제를 해결해야 할 필요성 자체가 사라지게 된다. 드레츠키는 이런 식으로 유비가 적용되지 않기를 원하지만, 이러한 생각은 외재적인 내용이 그 시점에서의 행동에 직접적으로 인과력을 행사하지 못한다는 사실 때문에 더욱 펼쳐내기가 힘들다.

이 대목에서 혹자는 드레츠키가 어째서 행동이 개인의 내적 상태에 의해서 충분히 결정되어 나온다는 전제를 고수하는지 의문을 품을 수도 있다. 앞서 예로 든 자판기의 비유에서 자판기가 동전의 내적 구성뿐만 아니라 사회적 관계에 의해 부여받는 교환가치까지 인식한다고 생각해 보자는 것이다. 물론 그런 일을 할 수 있는 자판기는 없겠지만, 인간은 실제로 그렇게 하고 있는 것처럼 보이지 않는가? 행동이 내적 상태에 의해 충분히 결정되어 있다는 전제를 포기한다면, 외재적 내용의 인과적 역할에 대해서

도 더욱 만족스러운 그림을 얻을 수 있을 것처럼 보인다. 2절에서 예로 든, C-섬유의 발화와 난로에서 손을 떼는 행동 사이의 관계에 대해 베넷이 말했던 바를 돌이켜 생각해 보자. C-섬유의 발화라는 한 물리적 사건 개별자는 복잡하고 비국지적인 배경 조건들이 함께 주어지는 한에서만 난로에서 손을 떼는 행동을 일으킨다. 베넷이 이 비국지적인 배경 조건들의 범위를 정확히 어디까지 생각하고 있는지는 모르겠지만, 심적 내용이 의존하고 있는 외적 환경에까지 범위가 넓어질 수 있다고 생각하더라도, 그녀의 글 자체에 이러한 생각에 반대할 것으로 보이는 구절은 없는 것 같다. 그러면 이제 내가 호랑이를 보고 피하는 행동을 하는 경우, 내 앞에 있는 그 호랑이가 지각과 판단을 매개로 나의 행동에 인과적으로 관련되어 있다고 생각할 수도 있다는 가능성이 보이게 된다.

드레츠키는 어쩌서 이렇게 쉬운 길을 놔두고, 외적 환경이 인과적으로 특정한 표상적 상태를 불러일으키고, 그 상태가 다시 과거에 완료된 학습의 결과 덕분에 행동을 일으킨다는 복잡하면서도 문제적인 설명을 제시하는 것일까? 심적 내용에 대해 반개체주의를 주장했듯이, 행동 설명에 대해서도 반개체주의를 주장할 수 있지 않을까? 문제는 이렇게 행동을 결정하는 데 고려해야 할 원인의 범위를 넓히는 일이 물리주의자의 입장에서 더 큰 부담을 발생시킨다는 것이다. 다시 자판기의 비유로 돌아가자면, 동전의 교환가치도 자판기의 ‘행동’을 결정하는 원인으로 취급하자는 말을 할 수는 있겠지만, 거기서의 인과 관계는 물리주의자들이 고려하고 있는 종류의 인과 관계가 아니다. 달리 말해서, 그것은 물리주의자의 입장에서 자연화될 필요가 있는 인과 관계이다. (드레츠키는 지각에 의해 표상을 가지는 것을 문제 없이 자연화된 인과 관계로 가정하고 있지만, 어쩌면 그 단계에서 이미 문제가 있었을지도 모른다.) 하지만 그런 작업이 여러 체계적인 이유들 때문에 만족스럽게 수행될 수 없기 때문에 비환원적 물리주의라는 입장이 나온 것이다. 결국 비환원적 물리주의자의 입장에서 남은 유일한 자연화의 수단은 심적 내용의 내용으로서의 인과적 역할을 이해하는 데에 별로 도움을 주지 못하는 총체적 수반 논제밖에 없게 될 텐데, 드레츠키는 여기에 만족하지 못하고 있는 것이다(Dretske 1989: 3).²⁶⁾

2.4. 결론

베넛의 작업들은 현재 비환원적 물리주의가 일반적으로 처해 있는 곤경을 잘 드러내 보여준다. 그것은 배제 문제를 해결하기 위해 물리적 속성에 대한 정신적 속성의 ‘의존’을 강조하는 방향으로 나아가는데, 그러면 두 속성이 정말로 구별되는지, 어떻게 구별되는지에 대한 의문이 남게 된다. 베넛과 같은 입장을 취하게 되면, 속성들의 개별화가 그것들 각각의 인과적 역할이 아닌 다른 무언가에 의해 이루어진다고 해야 할 것 같은데, 그것이 무엇인지는 알 수가 없다. 반대로 속성들의 개별화가 기존의 방식대로 인과적 역할에 따라 이루어진다고 한다면, 과잉 결정은 피할 수 없는 결론인 것처럼 보인다.

드레츠키는 비환원적 물리주의자의 입장에서 심적 내용이 내용으로서 가지는 인과적 역할을 더 잘 이해시키는 이론을 제공하려고 시도한다. 하지만 결과적으로 그의 이론은 많은 문제점들을 노출시키는데, 그 문제들은 물리주의라는 틀 내에서는 해결하기가 곤란한 것들이다. 그의 실패가 주는 교훈은 내용 외재주의를 받아들일 경우, 심적 내용의 내용으로서의 인과적 역할에 대해 총체적 수반 논제 이상의 것을 말하기가 어려우며, 그러면 다시 부수현상론의 위협에 시달리게 된다는 것이다.

비환원적 물리주의는 만족스러운 철학적 입장이 아니다. 그것은 정신적 속성의 인과적 효력에 대해 많은 사람들이 가지고 있는 생각을 충분히 반영하지 못하고 있다. 정말로 비환원적 물리주의, 혹은 그보다 더욱 강경한 물리주의에 만족하는 사람이 있다면, 그 사람은 나와 완전히 다른 세계관을 가지고 있는 사람이므로 설득이 불가능할지도 모른다. 하지만 내가 보기에 비환원적 물리주의를 주장하는 대부분의 논자들은 그것이 철학적으로 정당화될 수 있는 최대한 상식적인 입장이라고 생각하기 때문에 다소 불만족스러움을 느끼면서도 고수하는 것 같다. 이렇게 생각하는 사람들에게

26) 여기서 드레츠키가 특별히 총체적 수반을 고려하고 있지는 않다. 하지만 그가 말하고자 하는 바는 수반 논제만으로는 심적 내용의 인과적 역할을 제대로 포착할 수 없다는 것이므로, 총체적 수반 논제만 유지하려는 비환원적 물리주의에 대해서도 같은 식으로 생각했으리라고 충분히 예상할 수 있다.

대해서는 설득의 길이 열려 있다. 배제 문제를 일으키는 다섯 가지 주장들 가운데, 우리는 아직 물리 영역의 인과적 폐쇄성을 거부하는 대안에 대해 논하지 않았다.

3. 물리 영역의 인과적 폐쇄성 원리

이전 장에서 우리는 비환원적 물리주의가 부수현상론 비판에 대해서도, 정신적 속성이나 심적 내용이 가지는 인과력에 대해서도 만족스러운 답변을 제공해주지 못한다는 점을 보았다. 그렇다면 이제 주류적인 물리주의적 이론이 아니라 비물리주의적인 속성 이론론을 비롯한 여러 비주류적인 이론들에 눈을 돌려볼 때가 되었다.

각종 이론론, 다원론들은 정신적 속성이 물리적 속성과 확실하게 구별된다고 주장하며, 정신적 속성이 정신적 속성으로서 가지는 인과적 효력을 인정하는 데서도 아무런 어려움을 느끼지 않는다. 비주류적 이론들은 물리 영역의 인과적 폐쇄성 원리를 거부함으로써 정신적 속성의 확고한 인과적 효력을 유지하면서도 과잉 결정을 피해가고자 한다. 그런데 어떤 사람들에게 폐쇄성을 거부한다는 것은 용납할 수 없는 선택지인 것처럼 느껴지기도 한다.

인과적 폐쇄성의 원리는 주로 그것이 누린다고 생각되는 경험적 지지—즉, 현재 물리학의 이 원리에 부합하는 대단히 많은 사례들과 부합하지 않는 사례들의 결여라는 것의 결과로서 널리 받아들여지고 있다. 현재의 물리학은 많은 서로 다른 물리적 사건들을 위한 충분한 물리적 원인들을 발견했다고 논증된다. 더 나아가, 물리학은 물리적 사건을 위한 충분한 원인을 제공하기 위해 비물리적 원인에 호소할 필요가 결코 없었다. 이는 모든 물리적 사건이, 그것이 원인을 갖는 모든 시점에서 충분한 물리적 원인을 가진다는 주장을 위한 귀납적 증거를 제공하는 것으로 생각된다(Gibb 2013: 3).

이 원리를 포기하는 것은 원리적으로 물리적 현상에 대한 완전한 물리적 이론이 있을 수 없다는 것, 이론 물리학이, 그것이 완전한 이론이 되기를 열망하는 한에서, 순수한 물리학이기를 멈추고 환원 불가능하게 비-물리적인 인과력들—생기 원리, 엔텔레키, 정신적 에너지, 생의 약동, 혹은 정체를 모를 것에 호소해야만 한다는 것을 인정하는 것이다(Kim 1993b: 209).

위와 같은 생각을 가지고 있는 사람들이 보기에 물리 영역의 인과적 폐쇄성 원리는 현재 물리학의 이론과 실천에 의해 귀납적으로 정당화되는 원리이다. 그 원리를 거부한다는 것은 가능하기는 하겠지만 비합리적인 태도로 취급될 것이며, 심지어 물리학 내부에 초자연적인 원인을 도입해야 한다는 주장으로까지 받아들여지기도 한다. 불만족스러워하면서도 비환원적 물리주의를 고수하는 사람들이 많은 것도 은연중에 이러한 생각을 가지고 있기 때문이다.

이번 장에서 나는 현재 물리학의 이론과 실천은 물리 영역의 인과적 폐쇄성을 위한 귀납적 증거를 제공하기는커녕, 오히려 반대로 물리 영역의 인과적 개방성을 위한 귀납적 증거를 제공한다는 점을 보여주려고 할 것이다. 그러면 자연스럽게 어떻게 물리 영역이 개방적임에도 불구하고 비물리적 원인을 도입하지 않고 잘 실행되고 있는지가 드러날 것이다.

3.1. 근본 법칙의 조건부적 성격

물리 영역의 인과적 폐쇄성 원리는 많은 사람들에게 경험적 지지를 얻고 있다고 생각되고 있다. 수많은 물리학의 사례들이 그 원리를 입증하며, 그것을 반증하는 사례는 없다고 말해진다. 따라서 그 원리는 물리학이 따르는 확고한 방법론적 지침으로도 간주된다. 모든 물리학도는 물리적 현상을 설명하기 위해 오로지 물리적 원인만을 찾아 나선다. 물론 현재 물리학의 이론과 실천이 제공하는 것처럼 보이는 증거들로부터 물리 영역의 인과적 폐쇄성 원리를 주장하기 위해서는 약간의 도약이 필요하다. 폐쇄성 원리는 존재론적 주장이다. 그것은 모든 알려진 물리적 현상들이 충분한 물리적 원인에 의한 설명을 갖는다고 말하지 않고, 원인이 있는 모든 물리적 결과들이 충분한 물리적 원인을 갖는다고 말한다. 반면에 우리에게 주어져 있다고 가정되는 물리학의 증거들은 인식론적인 것이다. 그러므로 여기서 증거와 가설 사이의 관계는 좁은 의미에서 귀납적인 것이라기보다는

넓은 의미에서 귀납적인 것, 즉 가추적인 것이다. 물리 영역의 인과적 폐쇄성 원리는 현재 물리학의 관행을 가장 잘 설명하는 가설이라고 받아들여지고 있다. 현재 물리학의 관행은 넓은 의미에서 귀납적 증거를 제공하는 만큼, 폐쇄성 원리라는 가설을 결정적으로 입증했다고 수는 없겠지만, 귀납 추론 자체가 제기하는 철학적 문제들을 제쳐놓는다면, 그와 같은 증거들로부터 폐쇄성 원리를 주장하는 것은 상식적이고 합리적이라고 받아들여진다. 또, 폐쇄성 원리 자체가 물리주의를 함축하는 것은 아니지만(심신 평행론 등이 가능하므로), 어떤 물리주의자들은 이 원리의 성립이 물리주의를 믿을 만한 강력한 근거를 제공한다고 생각하기도 한다.

내가 보기에 위와 같은 일련의 추론은 증거 수집 단계에서 이미 문제가 있다. 그것은 주어진 증거를 잘못 해석하고 있다. 하지만 이 점을 보려면, 위와 같은 식의 생각들이 오랫동안 만연해 있었기 때문에 복잡한 단계들을 거쳐야 할 것이다.

우선 물리 영역의 폐쇄성 원리가 이해되는 직관적인 그림에 대해 생각해 보자. 그 그림에 따르면 물리적 현상은 다음과 같이 작동한다. 일단 폐쇄성 원리는 일반적으로 미시물리적 영역에서 성립한다고 생각된다. 하지만 여러 가지 이유에서 현재 우리가 가지고 있는 미시물리학은 최종적인 이론이 아니며, 언젠가 도래할 이상적인 미시물리학 이론에 의해 대체될 것이라고 많은 사람들이 믿고 있다. 어쨌든 그러한 최종적인 물리학 이론이 말하는 존재자들이 있다고 해 보자. 우리는 그 존재자들이 현재 물리학에서 말하는 것과 같은 입자들일지, 아니면 다른 어떤 종류의 존재자들일지, 감조차 잡지 못하고 있기 때문에 그냥 중립적으로 물리적 존재자들이라고 하는 수밖에 없을 것이다. 하지만 1장에서 보았듯이 데이빗슨을 비롯한 많은 사람들은 이상적인 물리학에 모종의 예외없는 엄격한 법칙만큼은 있을 것이라고 생각한다. 물론 우리는 이 근본적인 법칙이 몇 개나 될지, 어떤 내용을 가지고 있을지 전혀 예상하지 못하고 있지만, 이 법칙이 할 일에 대해서는 알고 있다고 여긴다. 엄격한 법칙은 근본적인 물리학 개념들 사이의 관계를 밝힘으로써, 물리적 존재자들의 행동을 완벽하게 기술하고 설명한다. 이것이 많은 사람들이 생각하는 법칙의 모습이다.

하지만 법칙이 위와 같은 일을 하고 있지 않으며, 할 수도 없다는 비판이 제기된 바 있다. 이미 오래 전에 앤스콤(Anscombe 1971/1993: 94)이 이런 비판의 단초를 제공했으며, 비교적 최근에는 카트라이트(Nancy Cartwright), 뉘프레(John Dupré), 기어리(Ronald Giere) 등의 철학자들이 위와 같은 식으로 이해된 법칙과 과학관에 지속적인 비판을 가해 왔다. 그 중에서도 특히 카트라이트는 앤스콤의 영향 하에,²⁷⁾ 물리 법칙에 대한 주류적인 이해를 비판하고 대안을 제시하는 데에 가장 많은 열정을 보였다. 이 글에서는 주로 카트라이트의 견해를 중심으로 비판을 전개해 나가도록 하겠다(Cartwright 1983, 1989, 1999).

법칙에 대한 카트라이트의 주장은 저작 시기와 맥락에 따라 크게 세 가지로 읽힐 수 있다.

- (1) 예외 없는 보편 양화 진술로 읽는다면, 법칙은 거짓이다.
- (2) 보편 양화 진술로서의 법칙은 참이지만, 오직 조건부로만 참이다.
- (3) 법칙은 역량(capacity)에 대한 서술로 읽는다면, 무조건적 참이다.

(1)은 주로 초기 저작에서 표명되는 입장이며, (3)은 중, 후기 저작에서 나타나기 시작하는 입장이다. (2)는 저작 시기를 막론하고 골고루 나타나는데, 시기에 따라 달라지는 주장들을 뒷받침하기 위해 다른 방식으로 이용된다. 비판적인 목적을 위해서는 일단 (1)과 (2)에 집중하는 것이 필요하다. 카트라이트가 초기 저작에서 선보인 유명한 예를 살펴보도록 하자(Cartwright 1983: 57ff).

잘 알려져 있듯이, 중력 법칙은 $F = Gmm'/r^2$ 라는 공식으로 표현된다. 이 법칙을 어떻게 이해할 것인가? 질량을 가진 ‘임의의’ 두 물체 사이에 작용하는 힘은 각각의 물체들의 질량의 곱에 비례하고, 물체들 사이의 거

27) 앤스콤의 영향에 대한 명시적 인정은 Cartwright(1999: 135)에서 찾아볼 수 있다.

리에 반비례한다고 해석한다면 거짓이다. 왜냐하면 어떤 물체들은 대전되어 있는데, 대전된 물체들 사이에서는 전기력이라는 다른 힘이 고려되어야 하기 때문이다. 대전된 두 물체들 사이에 대해서는 쿨롱의 법칙인, $F = k_e qq' / r^2$ 이 적용된다. 쿨롱의 법칙 역시 ‘임의의 대전된 두 물체들 사이에서 작용하는 힘은 각각의 물체들의 전하량의 곱에 비례하고, 물체들 사이의 거리의 제곱에 반비례한다’는 식으로 해석한다면 거짓이다. 왜냐하면 모든 물체는 전하량뿐만 아니라 질량 또한 가지기 때문이다. 따라서 질량을 가지고 있고, 대전된 두 물체들 사이에서 작용하는 힘을 알기 위해서는 중력 법칙과 쿨롱의 법칙이 상호작용하여 나오는 결과적 힘을 구해야 한다.

위의 사례로부터 카트라이트는 중력 법칙과 쿨롱의 법칙이 모두 거짓이라고 결론을 내린다. 물론 이는 법칙을 예외 없는 보편 양화 진술로 읽었을 경우이다. 거짓인 진술로 읽히고 싶지 않다면, 각각의 법칙들은 조건절을 첨가해야 한다. 이때 필요한 조건절은 ‘다른 힘이 작용하지 않는다면’이다. 그러면 중력 법칙은 ‘중력 이외의 다른 힘이 작용하지 않는다면, 질량을 가진 임의의 두 물체 사이에서 작용하는 힘은 각각의 물체들의 질량의 곱에 비례하고, 물체들 사이의 거리에 반비례한다’고 해석될 것이다. 이렇게 해석한다면, 중력 법칙은 참이다. 하지만 조건절이 붙은 법칙은 기존에 사람들이 기대했던 설명력을 가지지 못한다고 카트라이트는 지적한다.

그렇지만 그것은 매우 유용한 법칙은 아니다. 중력 법칙의 주요 업무들 가운데 하나는 다양한 복잡한 상황들 속에서 대상들이 경험하는 힘들을 설명하는 것을 돕는 것이다. 이 법칙은 오직 매우 간단하거나, 이상적인 상황들 속에서만 설명할 수 있다. 그것은 단지 중력만이 작용할 때 어째서 힘이 그러한지를 설명할 수 있다. 하지만 그것은 중력과 전기가 모두 문제가 되는 경우들에 도움이 되지 않는다. 일단 **다른 조건이 같다면**(*ceteris paribus*)이라는 수식어구가 붙으면, 중력 법칙은 보다 복잡하고 흥미로운 상황들에 무관하다(Cartwright 1983: 58).

이렇게 조건부 법칙은 참이더라도 복잡하고 구체적인 현실을 설명하기에는 역부족이다. 그것은 현실을 표상하지 않는다.²⁸⁾

이어서 카트라이트는 예상 가능한 문제 제기에 답변한다. 그 문제 제기에 따르면, 위의 논증에서 카트라이트는 중력 법칙이나 쿨롱의 법칙이 마치 합성력(resultant force)을 기술하는

것으로 생각했기 때문에 거짓이라고 잘못 판단을 내렸다. 하지만 각각의 법칙들이 분력(component force)만을 기술하는 것이라고 생각하고, 벡터 합에 의해 합성력을 산출한다면, 법칙들은 참이지 않을까? 초기 저작에서 이 물음에 대한 카트라이트의 답변은 복잡하고 미묘하다.

우선 힘들을 더하는 것은 자연이 아니라 우리가 하는 것이다. 자연에 분력은 존재하지 않는다. 그것은 우리가 계산을 편리하게 하기 위해 도입한 장치일 뿐이다. 이 답변은 기이하게 느껴질 수도 있다. 실제로 어떤 물체에 북쪽 방향으로 미는 힘과 동쪽 방향으로 미는 힘이 동시에 작용하는 상황은 충분히 있을 수 있다. 그때 물체는 북동쪽 방향으로 움직일 것이다. 카트라이트도 이러한 직관이 강력하다는 점은 인정한다. 하지만 법칙들이 분력을 기술하기 때문에 참이라는 식의 견해는 이 시절의 카트라이트에게 받아들여지지 않는데, 그 한 가지 이유는 그녀가 경험주의 전통에 충실했던 측면이 있기 때문이다.

각각의 법칙들이 참이라면 그것들이 기술하는 분력들은 실재할 것이다. 그 경우 합성력은 어떤 의미에서는 실재한다고 볼 수도 있지만, 벡터 합에 앞서 주어지는 분력들에 비해 파생적인 지위를 가지게 된다는 의미에서는 실재하지 않는다. 반면에 카트라이트의 경우에는 합성력이 일차적으로 실재하며, 분력은 편리한 허구에 지나지 않는다. 어느 쪽도 분력과 합성력이 모두 같은 수준에서 실재한다고 보지는 않는다. 그러면 벡터 합 견해는 무

28) 보다 정확하게 말하자면, 법칙의 이러한 비사실적 성격은 카트라이트가 ‘이론적 법칙’, ‘설명적 법칙’ 혹은 ‘근본 법칙’ 등으로 부르는 일군의 법칙들에만 적용되는 것이다. 그녀가 ‘현상학적 법칙’이나 ‘인과적 법칙’ 등으로 부르는 다른 종류의 법칙들은 현실을 표상하는 역할을 한다. 후자의 법칙들은 법칙이라는 이름으로 불리기는 하나, 보통 간단한 수식으로 표현되지 않으며, 심지어 다 쓰는 것이 불가능할 정도로 복잡할 수도 있다(Cartwright 1989: 225). 우리는 법칙이라고 하면 보통 물리학의 근본 법칙과 같은 것을 떠올리게 되며, 카트라이트가 표적으로 삼고 있는 것도 그러한 종류의 것이다.

엇이 문제인가? 우리는 지금까지 중력과 전기력을 논할 때 ‘힘(force)’이라는 말을 아무 문제가 없는 것처럼 사용했다. 하지만 경험주의적인 전통에서 볼 때, 힘은 그 자체로 경험 가능한 대상이 아니다. 경험, 혹은 관측 가능한 것은 물체의 가속도이다. 그래서 법칙은 마치 힘에 대해 쓰고 있는 것 같지만, 실제로는 물체들에 관한 보편 양화 진술로서 이해된다. 한 물체에 동쪽으로 미는 힘과 북쪽으로 미는 힘이 동시에 작용하여 물체가 북동쪽으로 움직이는 상황에서, 측정 가능한 벡터량은 물체가 북동쪽 방향으로 움직이는 가속도이지, 동쪽 방향의 가속도와 북쪽 방향의 가속도가 아니다. 따라서 분력들의 실재성을 인정한다는 것은 관측 불가능한 힘을 인정한다는 것이고, 이는 경험주의적인 전통에서는 부담스러운 일이다.

물론 힘을 측정하는 다른 방법이 있다면 이야기가 달라질 수도 있다. 하지만 일반적으로 말해서 물체에 가해지는 하나의 힘이 다른 하나의 힘에 의해 영향을 받고 상쇄되는 것이 가능하다면, 분력의 벡터 합에 의존하여 법칙의 사실성을 구제하려는 시도는 경험주의적인 전통과 조화되기 어렵다. 그러면 그러한 시도를 하는 쪽에서 경험주의적인 전통을 포기하면 어떻게 되는가? 그렇게 할 경우, 우리는 관측 불가능하지만 영향력을 발휘하는 힘의 실재성을 인정하는 것인데, 그것은 흥미롭게도 이후에 카트라이트가 발전시킨 바 있는 법칙에 대한 세 번째 해석, 즉 법칙을 물체들의 행동이 아닌 역량에 관한 진술로 해석하는 견해와 가까워지는 것이다. 어쨌든 이 시절의 카트라이트는 그러한 해석을 거부한다.

분력들의 실재성을 거부하는 것 말고도 카트라이트에게는 추가적인 답변이 있다. 법칙이 분력만을 기술한다 하더라도, 벡터 합은 합성력을 구하는 일반적인 방법이 될 수 없다는 것이 그것이다. 좀 더 일반적으로 말해서, 힘들의 합성을 주장하는 견해에서는 각각의 힘들의 효과를 기술하는 법칙들, 그리고 그 효과들 사이의 상호 작용을 구하는 법칙(벡터 합 법칙 등등)이 주어진다면, 최종적인 합성력을 구할 수 있을 것처럼 이야기한다. 하지만 카트라이트가 각종 사례들을 통해 보여주는 바는 대부분의 경우, 법칙들이 주어진더라도 그러한 여러 법칙들 사이의 상호 작용을 도출하는 법칙, 혹은 이론은 없으며, 있다고 하더라도 실제 일어나는 결과를 설명하

는 데는 별로 도움이 되지 않는다는 것이다(Cartwright 1983: 65-66). 그런 점에서 벡터 합이 유용하게 사용되는 전형적인 고전 역학의 사례들은 오히려 극히 예외적인 경우라고 볼 수 있다.²⁹⁾

법칙이 예외 없는 보편 양화 진술로서 참이라고 주장하고 싶어하는 논자들은 다른 전략을 취할 수도 있다. 그들은 중력 법칙이나 쿨롱의 법칙이 조건부로만 참이라는 사실을 받아들인다. 대신 그들은 각각의 하위 법칙들을 포섭하는 예외 없는 상위 법칙이 있다고 주장한다. 포섭된 하위 법칙은 포섭한 상위 법칙으로부터 연역적으로 도출된다. 하지만 방금 본 벡터 합 전략에 대한 비판은 논리 경험주의가 주장하는 포섭 법칙 모델에도 그대로 적용된다. 포섭 법칙이 있는 경우는 드물고, 있다고 하더라도 설명하는데 그렇게까지 도움이 되지는 않으며, 하위 법칙들의 도움 없이는 현상의 중요한 측면을 설명하지 못한다.

포섭 법칙 모델을 비판하기 위해 카트라이트가 들고 있는 예들 중 하나는 스넬의 법칙이다(Cartwright 1983: 46ff). 스넬의 법칙은 광학 법칙 가운데 하나로서, 매질의 굴절률과 입사각, 굴절각 사이의 관계에 관해 기술하는 법칙이다. 예컨대, 공기와 물처럼 굴절률이 서로 다른 두 매질이 맞닿아 있을 때 매질을 통과하는 빛의 경로는 매질마다 광속이 다르므로 휘어지게 된다. 이때 각각의 매질에서 빛이 휘 정도를 빛의 입사 평면 상에서의 각도인 θ_1 , θ_2 로 표시하고, 매질의 굴절률을 n_1 , n_2 라 한다면, 스넬의 법칙은 $\sin \theta_1 / \sin \theta_2 = n_2 / n_1$ 이라는 식으로 표현될 수 있다. 모든 매질 하에서 입사각과 굴절각 사이의 관계를 주장하는 보편 양화 진술로 본다면, 스넬의 법칙은 거짓이다. 왜냐하면 그것은 오직 등방성 매질에 한해서만 참이기 때문이다. 그러면 ‘오직 등방성 매질에 한해서만’이라는 조건절이 붙는 정제된 스넬의 법칙이 참이라고 해야 할 것이다. 하지만 대부분의 매질은 비등방성이기 때문에 정제된 스넬의 법칙은 많은 현상을 포섭하지 못한다. (비등방성 매질에서는 두 개의 굴절된 광선이 나타난다.) 게다가 우리는 비등방성 매질에서 성립하는 법칙을 알지 못하며, 앞으로 알 수 있

29) 따라서 전통적인 경험주의를 포기한다면, 법칙이 기술하는 분력들의 벡터 합에 의한 합성력의 산출은 드물지만 정당화될 수 있다는 뜻으로 읽히기도 한다.

우리라는 보장도 없다. 그럼에도 불구하고 우리는 비등방성 매질에서 빛의 굴절을 설명할 때 스넬의 법칙을 사용한다. 그 이유는 무엇인가?

우리는 같은 방식으로 비등방성 매질에서 두 굴절된 광선들의 각도들을 설명하기로 **결정한다**. 우리는 그 결정을 위한 좋은 이유들을 가지고 있을지도 모른다. 이 경우에 만약 매질이 거의 등방성이라면, 두 광선은 매우 가까이 있을 것이고, 스넬의 법칙이 예측한 각도에 가까울 것이다. 혹은 우리는 물리적 과정들의 연속성의 존재를 믿는다. 그러나 어쨌든 이 결정은 자연 법칙들에 대한 우리의 지식에 의해 강제되지 않는다(Cartwright 1983: 48).

카트라이트의 요점은 설명에서 법칙이 사용되기는 하지만, 논리 경험주의자들이 생각하는 것처럼 포섭에 의한 방식으로 사용되는 것은 아니라는 것이다. 그것은 우리가 전자기학의 법칙으로부터 스넬의 법칙을 유도할 수 있다는 사실을 안다고 하더라도 달라지지 않는다. 반면에 포섭 법칙 모델을 주장하는 사람들은 우리가 스넬의 법칙 같은 조건부 법칙을 사용하는 이유가 그것이 비등방성 매질에 적용되는 미래의 진정한 포섭 법칙으로부터 도출될 참된 결론에 충분히 가까운 결론을 낼 것으로 기대하기 때문이라고 말하고 싶어한다. 하지만 현실의 과학적 설명은 그러한 포섭 법칙에 대한 기대 없이도 이루어지고 있는 듯 보인다. 더욱 중요한 점은 설령 비등방성 매질에까지 적용되는 포섭 법칙이 있다고 하더라도, 우리는 거의 등방적인 매질에 대해 그것을 마치 등방적인 매질인 것처럼 간주하고 기존의 스넬의 법칙을 사용할 텐데, 그 결정은 포섭 법칙으로부터 나오는 것이 아니다. 그러면 우리는 어떻게 거짓이거나 조건부로만 참인 법칙을 통해 현상을 설명하는가? 카트라이트의 답은 우리가 법칙이 적용될 수 있도록 모델을 구성함으로써 그렇게 한다는 것이다. 이제 모델의 역할에 대해 좀 더 자세히 보도록 하자.

3.2. 모델을 통한 설명

이상의 논의를 통해 확인했듯이, 물체에 관한 보편 양화 진술로서 읽었을 때 근본 법칙의 조건부적 성격은 받아들이지 않기가 힘들다. 이제 우리가 법칙의 조건부적 성격을 받아들이는 데에 합의했다고 해 보자. 법칙의 조건부적 성격은 법칙의 설명 가능 범위에 제약을 가하는 것 같다. 물론 조건부 법칙이 유보 없이 참인 것처럼 보이는 이상적인 상황들이 있다. 카트라이트는 그런 상황들은 ‘법칙적 기계(nomologocial machine)’이라고 부른다.

법칙적 기계란 무엇인가? 그것은 올바른 종류의 (충분히) 안정적인 환경에서 반복된 동작으로 우리가 우리의 과학적 법칙들에서 표상하는 종류의 규칙적 행동들을 발생시킬, (충분히) 안정적인 역량들과 더불어 있는 구성 요소들, 혹은 요인들의 (충분히) 고정된 배열이다(Cartwright 1999: 50).

역량에 관한 이야기는 아직 하지 않았지만, 인과력이라고 바꿔서 읽으면 이해하기에 문제는 없다.³⁰⁾ 법칙은 법칙적 기계에 관한 서술로 읽는다면 참이다. 하지만 법칙적 기계, 혹은 이상적 상황과 그러한 상황에 충분히 가까운 상황은 그렇게 흔치 않다. 어떤 경우 그것은 자연에 주어져 있기도 하다. 대표적인 예로는 태양계가 있다. 고전 역학은 태양계라는 비교적 법칙적 기계에 가까운 상황이 자연에 주어져 있었기 때문에 놀라운 성공을 거둘 수 있었다.³¹⁾ 대부분의 경우 이상적 상황은 자연에 주어져 있지 않다. 그럼에도 불구하고 우리는 이상적 상황에서만 참인 조건부적 법칙을 가지고 이상적이지 않은 상황에 대해서도 설명을 하거나, 적어도 하려고 시도하는 것처럼 보인다. 어떻게 이런 일이 이루어질 수 있는가?

이 물음에 대한 카트라이트의 대답은 설명 혹은 이해는 규칙성을 전제하고, 규칙성은 법칙적 기계를 전제한다는 것이다(Cartwright 1999:

30) 흥미롭게도 전통적인 의미에서의 힘(power) 개념을 받아들이지 않는 흠직한 전통의 철학자들도 인과력(causal power)이라는 말이 무엇을 의미하는지 잘 알고 있으며, 심지어 자신들이 그 말을 특별한 설명 없이 사용하기까지 한다.

31) 앤스콤은 뉴턴 천문학의 성공이 어떤 의미에서는 지적인 재앙이었다고 말한다. 왜냐하면 그 성공은 과학적 설명의 본성에 관한 왜곡된 이해를 확산시켰기 때문이다(Anscombe 1971/1993: 99).

58-59). 설명이 필요 없을 정도로 자명하게 이해되는 현상(예: 지레)이 있더라도 거기에서 우리는 아주 단순한 법칙적 기계를 발견할 수 있다. 바꿔 말하자면, 우리가 어떻게든 비이상적 상황에서 법칙적 기계를 발견하거나 만들어내지 못한다면 설명은 주어질 수 없다는 뜻이다. 이는 상황이 정말로 어쩔 수 없을 정도로 비이상적이라면, ‘과학적’ 설명은 불가능하다는 뜻으로도 해석할 수 있다.

예컨대, 천 달러짜리 지폐를 광장에서 하늘로 던지는 상황을 생각해 보자(Cartwright 1999: 27).³²⁾ 천 달러짜리 지폐는 바람이 휘날리다가 광장 어딘가에 떨어질 것이다(물론 그렇지 않을 가능성도 있다). 지폐가 광장 어딘가에 떨어질 것이라고 예측하는 것도 어떤 의미에서는 설명일 수 있지만, 과학적 설명은 아니다. 과학적 설명으로 인정받으려면 그에 합당한 정밀성이 충족되어야 하기 때문이다. 과학이 이 상황에서 요구하는 정밀성의 한도 내에서 지폐가 언제 어디에 떨어질지 예측할 수 있어야 과학적 설명이라 부를 수 있다. 카트라이트는 이 경우 우리가 과학적 설명을 제공할 수 없다고 본다. 왜냐하면 이와 같은 상황에 적용할 수 있는 역학적 모델이 부재하기 때문이다. 보다 정확히 말하자면, 우리는 이 상황에 적용할 수 있는 부분적 모델만을 가지고 있다. 가령, 우리는 자유 낙하하는 물체로서의 지폐에 관한 모델은 가지고 있다. 하지만 지폐의 특정한 형태가 낙하에 미치는 효과에 대해서 생각하기 시작하면 벌써 자신이 없어진다. 바람이 미치는 효과에 대해서는 더욱 그러하다. 벡터 합을 통해 합성력을 구하고 싶어도 이 상황에 관련된 모든 분력들이 애초에 손에 잡히지 않는 것이다. 이런 식으로 우리가 흔히 물리적 현상이라고 분류하는 어떤 현상에 대해서조차 과학이 설명을 제공하지 못하는 경우는 매우 흔하다.

$F=ma$ 라는 역학의 가장 근본적인 법칙이 가지고 있는 표면적인 모습 덕분에, 많은 사람들은 질량을 가진 물체의 가속도에 영향을 미치는 물리적 힘은 어떤 것이든지 물리학에서 다룰 수 있다고 착각하기 쉽다. 하지만 가속도의 법칙이라는 내부 원리에서 힘은 질량이나 가속도와는 다른 추상 수준에 있는 개념이다. 추상적인 힘 개념은 실제 현상에 적용되기 위해 보

32) 이 예는 본래 오토 노이라트가 든 것이라고 한다.

다 구체적인 개념이 될 필요가 있다. 이 역할을 담당하는 것이 교량 원리이다. 교량 원리는 추상적인 개념이 보다 구체적인 상황에서 어떻게 적용되어야 할지를 알려준다. 자유 낙하하는 물체의 경우에 힘은 $-mg$ 가 될 것이고, 용수철의 경우에는 $-kx$ 일 것이고, 기타 등등. 교량 원리 없이 내부 원리만 있었다면, 고전 역학의 설명 가능 범위는 지금과 비교할 수도 없을 정도로 좁았을 것이다.

광장에서 뿌려진 지폐의 경우, 우리는 바람의 힘 등등의 요소들에 적용할 수 있는 교량 원리를 가지고 있지 않다. 교량 원리를 가지고 있지 않다는 말은 그 현상에 적용할 수 있는 모델을 가지고 있지 않다는 뜻이기도 하다. 그래서 우리는 문제가 되는 상황에서 오직 부분적 모델밖에 가지고 있지 않다고 말하는 것이다.

이상의 설명에서 교량 원리와 모델, 그리고 법칙적 기계 사이에 아주 긴밀한 관계가 있다는 사실이 드러났다. 정리하자면 다음과 같다. 우선 교량 원리는 가장 근본적인 법칙, 혹은 내부 원리가 보다 구체적인 상황에 어떻게 적용되어야 할지를 지시한다. 교량 원리가 주어지면, 우리는 모델을 구성할 수 있다. 카트라이트는 이러한 모델, 혹은 모델의 이러한 측면을 해석적(interpretive) 모델이라고 부른다. 해석적 모델은 구체적인 현실에 존재하는 사물들이 아니라, 이상화된 추상적 사물들을 다루고 있다. 추상적 사물들은 완전히 법칙에 따라 움직이도록 되어 있는데, 이는 달리 말하자면, 법칙이 현실이 아닌 모델을 기술하고 있다는 뜻이다. 법칙적 기계는 자연적으로든 인공적으로든 현실에 존재하는 구현된 모델이다. 따라서 모델은 법칙적 기계의 청사진이라고 할 수 있다.

이제 우리가 모델을 적용할 수 있는 곳에서만 과학적 설명을 기대할 수 있다는 사실을 인지했다고 해 보자. 그 다음 제기되는 물음은 우리가 어떻게 상황에 맞게 모델을 적용하느냐에 관한 것이다. 대부분의 상황들은 충분히 이상적이지 않으며, 따라서 모델을 그대로 적용하도록 잘 정돈되어 있지 않다. 그렇다면 현실을 모델에 맞게 재단하는 과정이 필요하다. 카트라이트는 우리가 모델을 매개로 이론을 현실에 적용하기 위한 첫 단계로서, ‘준비된 기술(prepared description)’을 마련한다고 말한다.

이론 진입의 첫 단계에서 우리는 기술을 준비한다. 우리는 이론 속으로 가져오도록 현상을 제시한다. 가장 명백한 필요는 이론이 방정식을 짚짓는 기술을 써 내리는 것이다. 하지만 방정식을 풀기 위해 우리는 어떤 경계 조건이 사용될 수 있는지, 어떤 근사 과정이 타당한지 등등을 알아야 한다. 그래서 준비된 기술은 이것들을 상세화하는 정보를 주어야만 한다(…) 이론 진입의 이 첫 단계는 비형식적이다. 보다 좋은 그리고 보다 나쁜 시도들이 있을 것이고, 상당량의 실천적 지혜들이 도움을 주겠지만, 이론의 어떤 원리들도 우리가 어떻게 기술을 기술을 준비해야 하는지 말해주지 않는다. 우리는 우리의 선행하는, 준비되지 않은 기술로부터 사실들을 취하고, 이론의 수학적 필요에 맞도록 그것들을 표현하는 올바른 방법이 무엇인지 교량 원리가 우리에게 말해주기를 기대하지 않는다. 이 단계에서 정확성에 대한 점검 기준은 우리가 이론 속에서 우리가 이론 바깥에서 아는 사실들을 얼마나 잘 재현했느냐가 아니라, 오직 최종적인 수학적 얼마나 처리가 성공적일 것이냐이다(Cartwright 1983: 133-34).

준비된 기술을 마련한다는 것은 결국 교량 원리가 제공하는 해석적 모델로 현실을 해석한다는 것이다. 인용문에서뿐만 아니라 카트라이트가 다른 곳들에서도 수도 없이 강조했듯이, 이론, 즉 교량 원리에 의해 구체화된 내부 원리를 적용하기 위해 현실을 해석하는 과정에서 이론은 어떠한 지침도 제공하지 않으며, 우리가 따를 수 있는 정형화된 규칙도 없다. 그것은 해석 일반이 그러한 것처럼 판단력을 요구하는 과정이다. 해석 과정에서 적절한 추상화와 이상화 과정이 필요하다는 것은 누구나 알고 있는 일반적인 지침이지만, 구체적인 상황 속에서 적절한 추상화와 이상화가 무엇인지는 그 상황에 있는 좋은 판단력을 지닌 주체만이 알 수 있다.³³⁾

어떻게든 현실을 해석해 내는 데 성공한다면, 그리고 성공하는 한에서만, 우리는 과학적으로 다루기에 적합한 모델을 얻는다. 이때의 모델은 특

33) 카트라이트는 추상화와 이상화를 구분한다. 간단히 말해서, 추상화는 존재하는 속성을 없애는 것이고, 이상화는 존재하는 속성을 다루기 편리한 속성으로 바꾸는 것이다. 그녀는 추상화가 이상화에 선행하며, 추상화는 역량에 관한 지식을 전제한다고 본다(Cartwright 1989: Ch.5).

정 상황을 과학적으로 다루기에 적합한 방식으로 표상하는 모델이며, 그렇기 때문에 ‘표상적(representative) 모델’이라고도 불린다(Cartwright 1999: 180). 표상적 모델은 특정 상황을 표상하는 모델이므로, 앞서 말했듯이 이론에 의거해서만 구성될 수 없으며, 엔지니어링을 비롯한 여러 요소들에 의한 복합적 산물이다. 표상적 모델은 우리가 관심을 기울이는 어떤 중요한 측면에서 그것이 표상하는 현실과 유사함으로써 모델로서의 역할을 수행한다.³⁴⁾ 이는 우리가 목적에 따라 같은 현실에 대해서도 서로 다른, 심지어 서로 비일관적이기까지 한 모델들 가운데 하나를 구성할 수 있다는 뜻이기도 하다. 어쨌든 표상적 모델은 과학이 직접 다루는 가장 구체적인 수준이라고 할 수 있다.

법칙이 모델과 그것을 구현한 법칙적 기계에 대해서만 참이고, 물리 이론이 모델을 매개로 해서만 현실을 다룰 수 있다는 사실을 깨달으면, 물리 영역의 인과적 폐쇄성이라는 원리가 어째서 경험적 지지를 받지 못하는지, 그럼에도 불구하고 마치 경험적 지지를 받는 것처럼 보이는 가상이 나타나는지가 보이기 시작한다.

전형적인 당구공 물리학의 사례를 생각해 보자. 당구대 위에서 펼쳐지는 당구공들의 움직임은 우리가 일상에서 발견할 수 있는 법칙적 기계의 대표적 사례일 것이다. 그것은 고전 역학의 모델을 모범적으로 구현하고 있다. 이제 어떤 사람이 실제로 당구를 치는데, 우리는 그가 공을 친 뒤에 어떤 결과가 나올지 예측하고 싶어 한다고 해 보자. 그 상황에서 우리가 필요한 정보들을 모두 알고 있다면, 충분히 신뢰할만한 예측을 할 수 있다는 것이 일반적인 생각이다. 하지만 여기서 필요한 정보들을 모두 알고 있다는 것이 정확히 무엇을 의미하는가? 만약 지나가던 누군가가 갑자기 당구대를 흔들거나, 당구공을 집어서 도망간다면, 당구공의 움직임은 역학 이론의 예측으로부터 벗어나게 될 것이다. 우리는 그런 간섭을 미리 알지 못한다. 오히려 우리가 알고 있는 것은, 정밀한 예측이 이루어지려면 그런 간섭이 미리 배제되어야 한다는 것뿐이다. 다른 말로 하자면, 법칙적 기계는 제대로 작동하기 위해 그것의 작동을 방해할 수 있는 외부의 인과적

34) 이 점은 기어리(Giere 2008: 128)가 더욱 분명하게 강조하고 있다.

요인들로부터의 ‘차폐(shielding)’를 필요로 한다(Cartwright 1999: 57-58).

차폐가 제대로 이루어지지 않았다면, 남은 가능성은 두 가지밖에 없다. 우선 이제까지의 논의를 통해 보다 명백해진 한 가지는 그런 경우에는 과학적 설명이 주어질 수 없다는 것이다. 우리는 그런 변칙적인 상황에 적용할 수 있는 모델을 가지고 있지 않거나, 부분적 모델만을 가지고 있기 때문이다. 다른 한 가지는 그런 외적 요인의 간섭이 어떻게든 역학에서 말하는 힘 개념으로 재기술되는 경우이다(Cartwright 1989: 210). 어떤 유형의 간섭은 어쩌면 당구공들의 최종 상태를 예측하기 위해 고려해야 하는 힘들 중 하나로 취급되는 데에 성공할 수 있을지도 모른다. 그 경우, 우리는 그 특정한 외적 힘까지 구성 요소들 가운데 하나로 포함하는 새로운 법칙적 기계를 구성하는 것이다.

방금 제시한 사례는 현재의 물리학이 일반적으로 정신적 원인을 어떤 식으로 다루고 있는지를 알게 해 준다. 정신적 원인은 과학적 설명에서 두 가지 방식으로 취급되는 것처럼 보인다. 1) 차폐가 성공적으로 이루어질 수 있다면, 그것은 모델이 표상해야 하는 원인들 중 하나로 취급되지 않는다. 2) 모델 속에서 표상할 수 있는 힘들 중 하나로 해석될 수 있다면, 정신적 원인은 마치 물리적 원인들 중 한 가지처럼 취급된다. 1)과 2)가 모두 불가능하다면, 정신적 원인에 의한 물리적 결과의 발생은 물리학이 설명하려고 시도하는 현상이 아니다. 물리학이 정신적 원인을 비롯한 다른 원인들을 이러한 식으로 다루기 때문에 물리학이 비물리적 원인을 도입하지 않고도 잘 작동하는 것이다.

요컨대, 물리학은 자신이 설명을 할 수 있는 곳에서만 설명을 하고, 또 하려고 시도한다. 설명을 하는 경우에도, 그것은 여러 종류의 원인들을 모델 속에서 마치 한 종류의 원인들인 것처럼 다룰 수밖에 없다. 물리 영역의 인과적 폐쇄성 원리라는 가상은 현재 물리학의 설명 범위가 제한적이고, 그 제한적인 설명 범위 안에서도 특정한 설명 방식 취할 수밖에 없다는 점을 보지 못하기 때문에 생겨난 것이다. 모델 구성이 완료된 이후의 단계들에만 시선을 집중하면, 이 가상은 결코 깨지지 않을 것이다.

근본 법칙에 대한 위와 같은 논의들이 물리 영역의 인과적 폐쇄성을 고수하는 사람들에게 어떻게 받아들여질 것인가? 그들은 현재 물리학의 법칙에 대한 기술로서는 카트라이트가 제시하는 설명이 정확하다는 점을 충분히 받아들일 수 있다. 예컨대, 카트라이트와 상반되는 견해를 가지고 있는 호퍼는 현재 물리학에 대한 기술로서는 카트라이트의 반토대주의적 그림이 더욱 정확하다는 점을 인정한다(Hoefer 2008: 320-21). 하지만 그에게는 현재 물리학이 가지고 있는 이러한 특성들, 즉 법칙의 조건부적 성격과 이론 내적 지식만으로는 불가능한 모델 구성의 필요성 등등이 단지 현재 물리학의 결함으로만 보인다. 이 결함들은 최종적인 물리학에서는 극복될 것으로 그는 기대한다(Hoefer 2008: 318-19). 이러한 생각은 그만의 것이 아니다. 앞에서 보았듯이 데이빗슨도 “격을 갖춘 법칙이라면, 그 법칙을 적용하는 데 필요한 조건이 갖춰졌는지 아닌지를 우리가 미리 결정할 수 있도록 해줘야 한다”(데이빗슨 1974/2012d: 382)고 말함으로써 이상적인 물리학에 대한 유사한 생각을 표현한 바 있다.

이처럼 미래의 물리학에 기대를 거는 것은 한편으로는 이해할 만한 반응이다. 다음의 인용문을 보자.

나는 이러한 장치들[레이저와 초전도체]을 이해하기 위해 양자적 고려들이 얼마나 결정적인지에 감명받았지만, 또한 그것들[양자적 고려들]이 그것들 자체로는 얼마나 성취할 수 있는 것이 거의 없는지에도 마찬가지로 감명받았다. 그것들은 정확한 예측이 기대될 수 있기 이전에 대량의 고전 물리학들, 실천적 정보들, 소재들에 대한 지식, 그리고 마지막으로 대단히 조심스럽고 현명한 엔지니어링과 결합될 필요가 있다. 이것들 중 어떤 것도 양자 물리학의 언어로 기술되어 있거나, 혹은 원리적으로라도 기술될 수 있을 것처럼 보이지 않는다. 물리학의 자기-폐쇄성에 대한 나의 주저를 이끈 것은 이러한 연구들이었다(Cartwright 2010: 613).

여기서 카트라이트가 말하고자 하는 바는 양자 역학과 같은 현재의 첨단 분야를 이해하고 검증하기 위해서도 ‘과거의’ 이론들, 전이론적 혹은 이론 독립적 지식들이 필요한데, 이런 것들은 원리적으로라도 현재의 물리

학만으로 기술되지 않을 것처럼 보인다는 것이다. 다시 말해서 최신의 물리학만 가지고는 그것이 적용될 것으로 기대되는 미시 물리적인 현상들조차 다 포괄하지 못한다. 미시 물리적인 현상들을 설명하기 위해 미시 물리적이 아닌 원인들을 언급하는 것이 인과적 설명에서 필수 불가결하다면, 미시 물리적 수준에서의 인과적 폐쇄성은 성립하지 않는다. 현재 물리학의 관행으로부터 세계상을 추론해야 한다면, 자기-폐쇄적인 물리학이 토대를 이루고 있는 그림이 아니라 반토대주의적이고 다원론적인 그림이 훨씬 더 그럴듯하다는 것이 카트라이트가 초기부터 변함없이 드러내고 있는 생각이다. 그녀와 상반되는 입장을 가지고 있는 논자들이 보기에 이러한 생각은 거부하기가 어렵다. 특히 모델의 매개적 역할에 대해 민감하게 의식하고 있는 호퍼 같은 사람들은 현실에서의 물리학적 설명이 미시 물리학의 언어로 원리적으로라도 기술될 수 없는 것처럼 보이는 이유가 단순히 수학적 계산의 복잡성 때문이 아니라는 점을 잘 알고 있다. 그것은 모델을 구성할 수 있는 가능성조차 없어 보인다는 점에서 비롯된다. 그렇기 때문에 미래의 물리학이 어디에나 적용 가능한 모델과 그 구성 규칙까지도 제공해 줄 것으로 기대하는 것이다. 이는 달리 말하자면, 미래의 이상적인 물리학에 대한 기대가 불가능하다면 물리학의 폐쇄성은 경험적 지지를 받는다고 말할 수 없다는 뜻이다.

하지만 우리에게 언젠가 그런 이상적인 이론이 주어질 것이라고 믿을 만한 경험적 근거가 있는가? 호퍼와 같은 입장을 공유하는 논자들이 제시할 수 있는 가장 강력한 근거는 아마도 물리학의 이론이 점점 더 많은 현상을 설명할 수 있게 된 성공의 역사일 것이다. 하지만 광장에서 지폐를 던지는 사례에서 보았듯이 우리에게서는 무수한 실패의 역사 또한 존재하며, 카트라이트가 말하였듯이 성공의 역사와 실패의 역사 중에서 어느 쪽이 무거운지 측정할 수 있는 방법은 없다(Cartwright 2010: 618). 그러면 꼭 경험적 증거가 아니더라도 언젠가 이상적인 이론이 도래할 것이라고 믿을 만한 좋은 이유가 있는가? 어쩌면 자연의 단순성에 대한 믿음 같은 것도 이유가 될지는 모르겠다. 그러나 내가 카트라이트의 작업들을 통해 배운 바에 의하면, 그러한 이상적인 이론을 기대하기가 어려운 더욱 그럴듯한

이유가 있다. 이상적인 이론에 대한 기대는 잘 들여다보면 양립하기 어려운 기대들이 충돌하고 있는 것처럼 보인다. 지금 우리가 생각하는 대로라면, 미래의 근본 법칙 역시 현재의 그것과 마찬가지로 고도로 추상적이고 정밀하면서도 간결한 수식으로 되어 있을 것이다. 고도의 정밀성을 요구하는 개념으로 되어 있는 법칙은 그만큼 적용 범위도 제한적일 것이고, 외적 조건의 영향에도 민감할 텐데, 어떻게 조건부적이지 않을 수 있을까? 또, 간결한 법칙이 어떻게 현실에 적용하기 위한 완벽한 규칙들을 포함하고 있을 수 있을까? 복잡하고 구체적인 현실들을 표상한다는 목적과 근본 법칙의 엄격한 성격 사이에서는 해소 불가능한 긴장이 발생하는 것이 당연하게 여겨진다. 이는 우리의 언어와 사고 체계의 본성에 새겨져 있는 긴장이다. 미래의 이론에서 이 긴장이 완전히 사라지리라고 기대하는 것은 단순한 경험적 발견을 기대하는 것이 아니라 우리가 완전히 다른 존재가 되리라고 기대하는 것이다. 그러한 기대가 충족된다는 것이 무엇인지 지금 우리는 이해할 수 없다.

3.3. 조건부 법칙에서 역량에 관한 서술로

물리 영역의 인과적 폐쇄성 원리라는 가상을 해제하기 위해서는 근본 법칙의 조건부적 성격과 이론이 모델을 매개로 함으로써만 현실을 다룰 수 있다는 점을 강조하는 것으로도 충분하다. 하지만 어느 시점에서 카트 라이트는 법칙에 대한 자신의 초기 해석을 수정하여, 법칙이 사물의 역량에 관한 서술로 해석되는 것이 가장 그럴듯하다고 주장한다. 이제 역량은 법칙적 진술의 진리 확정자 역할을 한다.³⁵⁾ 여기서 역량은 힘(power), 본성/본질(nature), 실체적 형상 같은 아리스토텔레스주의적인 개념들과 거의 교환 가능한 개념으로 쓰이고 있다.³⁶⁾ 반면에 그녀는 흔히 유사하다고

35) 프루스(Pruss forthcoming)가 이런 식으로 아리스토텔레스의 형상을 법칙의 진리 확정자로 보고 있다.

36) 카트 라이트는 역량이 정확히 어떤 것이라고 말하기보다는 관련된 개념들과의 비교, 대조를 통해 보여주려고 하고 있다. 역량 개념이 직접 설명하기에는 너무 기초

생각되는 성향과는 이 개념들을 구분하고 있는데, 성향은 보통 그것의 소유와 표현 사이에 일대일 대응 관계가 성립한다고 생각되는 경향이 있기 때문이다. 반면에 한 역량의 표현은 매우 다양할 수 있다(Cartwright 1999: 59-64).

이 대목에서 카트라이트가 말하고자 하는 것은 다음과 같다. 예컨대, 소금은 물에 녹는 성향을 가지고 있는데, 만약 소금이 어떤 이유에서든 녹지 않는다면 물에 녹는 성향이 표현되지 않은 것이라고 흔히 생각된다. 반면에 북쪽으로 미는 힘(force)은 어떤 물체의 북쪽 방향 가속도를 증가시키는 역량을 가지고 있는데, 남쪽으로 미는 더 큰 힘이 가해져서 상쇄되면 물체의 북쪽 방향 가속도는 감소하겠지만, 그럼에도 불구하고 북쪽으로 미는 힘은 현실에서 작용하고 있다는 것이다. 원인과 배경 조건을 구분하지 않는다면, 그 상황에서 북쪽 방향으로 미는 힘은 해당 물체가 남쪽 방향으로 가속되는 원인들 중 하나이다. 카트라이트 자신이 들고 있는 사례는 더욱 좋다. 같은 부호로 대전된 입자들은 서로를 밀어내는 역량을 가지고 있다. 하지만 특수한 상황(Towfic Shomar가 설계한 장치) 속에서 같은 부호로 대전된 입자들은 그러한 역량을 가지고 있기 때문에 가까워질 수도 있다. 이 경우는 원인과 배경 조건을 직관적으로 구분하더라도 같은 부호로 대전된 것이 두 입자를 서로 가깝게 만드는 원인이라고 말할 수밖에 없을 것 같다. 또, 카트라이트는 세금이 가지고 있는 물가를 올리는 역량 같은 사회적 역량까지 고려하고 있는데, 사회적 역량이 상황에 따라 매우 다양한 표현들을 가진다는 점은 분명하다. 이런 이유에서 그녀는 역량을 일반적으로 생각되는 성향과 동일시하지 않는다. 그래도 이미 널리 사용되고 있는 성향이라는 개념을 통해 역량이라는 개념을 이해하고 싶다면, 특수한 성향이 아니라 고도로 유적인 성향과 동일시되어야 한다고 주장한

적이라고 생각하기 때문일지도 모른다. 역량의 존재론을 주장하는 철학자들 가운데 한 사람인 머포드도 역량, 혹은 힘이 무엇이라고 설명하기보다는 다른 개념들과의 관련을 통해 드러내려고 하고 있다. 특히 그는 모든 속성이 곧 역량이거나 역량들의 덩어리라고 말하고 있다(Mumford and Anjum 2011: 3-4). 또, 그는 어떤 경우에 역량들의 덩어리는 그것을 이루는 부분적인 역량들의 단순한 합 이상의 것일 수 있다고 말함으로써 창발적 역량에 대해서도 말하고 있다(Mumford and Anjum 2011: 104-105).

다.³⁷⁾

정리하자면, 하나의 역량은 보통 하나의 전형적인 표현과 연결되어 있지만 상황에 따라서는 다양하게 표현을 가질 수 있다. 중요한 것은 그 역량이 전형적으로 표현되지 않는 경우에도, 그것은 현실에서 작용하고 있다는 것이다. 그래서 역량은 그것과 연결된 전형적인 행동을 표현하려는 ‘경향성’을 가진다고도 말해진다(Cartwright 1999: 29).

카트라이트는 역량 개념을 단지 설명적 개념으로만 여기고 있지 않다. 그녀는 다양한 역량들이 완전히 실재적이라고 본다. 이에 대해 두 가지 물음이 제기될 수 있다. 어째서 법칙을 역량에 관한 서술로 읽는 것이 가장 그럴듯한가? 역량들이 실재적이라고 보아야 할 이유가 무엇인가?

우선 카트라이트는 초기 저작에서 이미 법칙을 인과력에 관한 서술로 읽을 수 있는 가능성을 고려한 적 있다. 1절에서 보았듯이, 초기에 카트라이트는 그러한 해석을 거부했는데, 그녀의 논증은 다음과 같은 것이었다.

만약 법칙이 물체의 행동에 관한 것이 아니라 물체가 가지는 인과력, 즉 역량에 관한 것이라면, 법칙은 각각의 분력들을 기술하는 참인 진술이다. 그렇다면 분력들은 실재적이다. 하지만 관측 가능한 것은 합성력의 효과뿐이므로 분력들의 효과는 관측 불가능한 속성이고, 관측 불가능한 속성은 시험할 수 없으므로, 분력들은 비실재적이다. 따라서 법칙은 역량에 관한 서술이 아니다.

변화된 카트라이트의 견해에서 핵심은 시험 가능한 것은 관측 가능한(occurrent) 속성만이 아니라는 것이다(Cartwright 1999: 65). 우리는 관측 불가능한 속성 또한 시험하고, 이해한다. 역량은 다양한 표현들을 가질 수 있고, 그 표현들 사이에 어떠한 공통적인 관측 가능한 속성도 없을 수 있지만, 그렇더라도 우리는 그 표현들이 한 역량으로부터 나온 것이라는 사실을 경험하고 안다. 따라서 법칙은 물체의 관측 가능한 행동이 아닌 역량에 관한 서술이다. 물론 이러한 입장 변화는 벡터 합에 의한 합성력의 산출이라는 방법에 대한 태도의 변화와도 함께 가는 것이다.

관측 불가능한 속성이 알려지고, 인과적 역할을 한다는 것은 법칙적 기

37) 내가 아는 한 최고로 유적인 성향은 인간의 자유이다. 자유라는 역량은 인간의 모든 행위로 표현된다.

계의 작동을 이해하는 데도 중요하다. 법칙적 기계는 제대로 작동하기 위해 그것을 방해할 수 있는 외부의 인과적 요인으로부터의 차폐를 전제하는데, 차폐하는 데 성공한 외부의 인과적 요인들은 관측 가능하지 않지만 인과적 영향력을 행사하고 있는 것이 분명한 힘들이기 때문이다.

요약하자면, 카트라이트는 우리가 역량 개념을 전제하고, 법칙을 역량에 관한 서술로 읽지 않으면 현재 과학의 관행을 설명할 수 없다고 주장하는 것이다.

그 다음, 역량의 실재성에 관한 물음에 대해서는 앤스콤(Anscombe 1971/1993: 93)을 연상시키는 논변이 주어진다. 역량의 실재성을 의심하는 논자들은 역량 개념이 편리한 허구라는 점은 인정하지만, 이 세계에 역량이라는 것이 존재하지 않더라도 세계에 관한 우리의 경험은 그대로일 것이라고 주장한다. 카트라이트는 이를 인과의 비실재성을 주장하는 흄의 회의주의와 비슷한 것으로 보고, 일상적 사물에 대한 관찰 경험이 이미 역량-적재적이라는 식의 대응을 하고 있다(Cartwright 1999: 69).³⁸⁾

변화된 견해는 다음과 같은 식으로도 이해될 수 있다. 역량 개념을 도입하기 이전의 카트라이트에게 우리가 어째서 모델을 통해 설명을 하느냐고 묻는다면, 아마도 도구주의적이고 실용주의적인 답변이 주어질 것이다. 반면에 역량 개념을 도입한 이후의 그녀는 현실에서 이미 작용하고 있는 다양한 역량들로부터 우리가 관심을 기울이는 역량의 효과만을 추출하기 위해 모델과 법칙적 기계가 필요하다고 응답할 것이다.

온갖 다양한 역량들이 실재하는 ‘얼룩덜룩한 세계(dappled world)’라는 존재론을 어떻게 평가해야 할까?

이에 대해 카트라이트와 상반되는 견해를 지닌 철학자가 비판적이라 하는 것은 쉽게 예상할 수 있지만, 의외로 그녀와 거의 유사한 과학 철학적 입장을 공유하는 기어리(Giere 2008)조차도 역량들의 존재론에 대해서는 비판적이다. 기어리는 카트라이트가 가지고 있는 다원론적 세계관을 과학의 방법론으로서만 받아들이고 싶어한다. 그것은 그가 물리주의 같은 다른 형이상학을 가지고 있기 때문이 아니라, 과학은 어떤 종류의 형이상학도

38) 이 글의 1장 3절 참조.

없이 수행될 수 있다고 믿기 때문이다. 다만 여기서 기어리는 마치 과학자들이 특정한 형이상학을 가지지 않아도 되는 것처럼 이야기하는데(Giere 2008: 131), 그것은 너무도 명백하다. 과학자들은 심지어 특정한 방법론조차 가지지 않아도 되는데, 이는 과학자 자신은 그가 어떤 방법으로 과학적 작업을 수행하는지에 대한 반성적인 지식을 가지지 않아도 된다는 뜻이다. 그에게 필요한 것은 오히려 구체적인 상황 속에서 어떻게 작업을 해 나가야 하는지, 무엇이 적절한 과학적 작업인지 아닌지에 대한 노하우일 것이다.

물론 과학자들은 지적으로 고립된 존재들이 아니며, 자연적 이유에서든, 다른 이유에서든 특정한 형이상학을 가지고 있을 가능성이 높다. 현대에는 특히 물리주의적인 형이상학을 가지기가 쉬울 것이다. 그리고 그런 형이상학은 토대주의적인 방법론을 제시하는 것처럼 보일 수도 있다. 토대주의적인 방법론에 따라 과학을 수행하는 것은 많은 경우 해롭다. 이런 식으로 만약 한 과학자가 가지고 있는 특정한 형이상학이 그가 과학적 작업을 수행하는 데 부정적인 영향을 미친다면, 그는 자신의 방법론을 교정할 필요가 있을 것이다. 하지만 그 경우에도 그에게 필요한 것은 다른 방법론을 가지는 것이 아니라, 그가 가지고 있는 해로운 방법론을 망각하는 것이다. 물론 특정한 형이상학과 방법론에 대한 반성적 의식을 가지게 된 사람에게 단순히 그것을 망각하라고 하는 것은 현실적으로 쉽지 않다. 그런 경우에는 교육적인 목적을 위해 다원론적 형이상학을 배움으로써 반토대주의적인 방법론 또한 가능하다는 것을 깨닫는 것이 유용할 수도 있을 것이다.

기어리가 정말로 하고 싶었던 말은 아마도 적절한 과학적 방법론(이는 일차적으로 과학 철학자들을 위한 것이다)이 특정한 형이상학 없이도 확립될 수 있다는 것이었으리라고 생각한다. 그는 건전한 반토대주의적인 방법론이 역량을 도입하지 않고도 ‘인과적 구조’에 의해 충분히 얻어질 수 있다고 주장한다. 여기서 그는 인과적 구조가 무엇이라고 정확히 밝히고 있지는 않지만, 문맥상 특정 상황에서의 인과 관계들의 배열을 의미하는 듯하다(Giere 2008: 131). 그런데 우리는 한 상황에 적용되는 과학적 설명이 유사한 다른 상황에서도 적용될 수 있기를 기대한다. 이론의 역할은 그러

한 일반화를 가능하게 하는 것이다. 심지어 우리는 같은 부호로 대전된 두 입자가 서로를 가까워지게 하는 상황에서도 쿨롱의 법칙을 이용해 두 입자의 움직임을 계산하려 한다. 그 말은 쿨롱의 법칙이 대전된 입자의 역량, 즉 어떤 상황에서도 같은 전하를 가진 다른 입자를 밀어내려는 경향성을 서술하고 있다고 보아야 한다는 뜻이 아닌가? 다른 무슨 대안이 있느냐고 카트라이트는 반문한다(Cartwright 2008).

역량 개념에 대한 기어리의 반감은 사실 과학사와 철학사에 대한 그의 이해와 연결되어 있다. 그는 과학 혁명과 계몽주의가 아리스토텔레스적인 세계관을 완전히 과거의 것으로 만들어버렸다고 생각한다. 반면에 카트라이트는 베이컨과 뉴턴 이후 오늘날에 이르기까지 우리는 우리 자신이 표면적으로 말하고 생각하는 것보다 훨씬 더 아리스토텔레스주의적이라고 주장한다. 내가 보기에 역량 개념 그 자체에 시대착오적인 것은 아무것도 없다. 오히려 기어리 자신도 은밀하게 역량과 친족적인 개념을 도입하고 있는 것처럼 보이는 구절도 있다.

중력에 관한 고전적 설명에서 일반 상대성 이론적 설명으로의 이행은 멋진 사례를 제공한다. 고전 역학의 맥락에서 카트라이트는 질량이 다른 질량들을 끌어당기는 역량을 가지고 있다고 말한다. 일반 상대성의 맥락에서, 그녀는 아마 질량들이 그것들 주위의 시공간을 왜곡시켜 다른 물체들이 자연스럽게 왜곡된 시공간에 있는 측지선으로 움직인다고(그리고 따라서 단지 끌어당기는 것처럼 보인다고) 말할 것이다. 그래서 우리는 질량들이 가지고 있는 역량에 관한 설명을 변화시킨 것이다. 그러나 여기서 나에게 역량에 관한 말하기는 명백히 잉여적인 것으로 보인다. 우리가 질량들 사이의 상호작용의 **인과적 본성(causal nature)**에 관한 우리의 이해를 원거리 작용에서 시공간의 유도된 구조에 의해 매개된 작용으로 변화시켰다고 말하는 것으로 충분하게 보인다(Giere 2008: 132-33, n.12, 강조는 인용자의 것임).

위의 인용문은 역량 개념을 도입하지 않고도 질량의 ‘작용(action)’에 대한 이해의 변화가 서술될 수 있다는 점을 보이려고 의도하고 있다. 하지만 그 과정에서 그는 이미 “인과적 본성”이라는 역량과 교환 가능한 개념

을 사용해버렸다. 그는 이 표현을 다른 표현으로 대체하려 하거나, 아니면 역량과 연결된 의미에서 사용한 것이 아니라고 주장할 것이다. 하지만 그의 시도들은 실패할 수밖에 없다. 질량이 다양한 상황들 속에서 공통적으로 하는 작용에 관해 말하기 시작하면, 이미 질량의 역량에 관해 말하고 있는 것이기 때문이다.³⁹⁾

기어리의 입장에 대해 마지막으로 한 가지만 더 지적하자면, 적절한 과학적 방법론이 특정한 형이상학을 전제하지 않고도 얻어질 수 있다는 그의 주장은 우리가 형이상학을 어떤 식으로 생각하느냐에 따라 다르게 평가할 수 있다. 형이상학이 과거의 제일철학과 같은 토대적 역할을 하는 것으로 이해된다면 기어리의 주장은 적절하다. 반면에 형이상학은 과학적 방법론과 분리되지 않는 것으로 이해될 수도 있다. 과학자들은 이미 법칙, 인과, 힘 같은 개념들을 사용하고 있는데, 적절한 과학적 방법론은 그와 같은 개념들이 어떻게 이해되고 사용되어야 하는지를 말해줄 것으로 기대된다. 그러한 작업은 형이상학과 선명하게 구분되기가 어렵다. 이 경우 형이상학은 제일철학이 아니라 오히려 과학적 방법론의 일부이면서 동시에 과학과 과학적 방법론에 대한 이해를 요약하고 일반화한 결론으로 보아야 할 것이다. 이는 반토대주의적 방법론이 철학에도 적용되고 있다는 것을 의미한다. 형이상학을 이런 식으로 이해한다면, 기어리의 주장은 요점을 상실한다.

기어리의 비판 속에서는 역량들의 존재론이 어째서 문제인지 잘 드러나지 않는다. 그는 이미 카트라이트와 너무도 많은 것을 공유하고 있다. 흄 같은 경험주의자라면 역량을 비롯한 각종 인과적 개념들의 도입 자체에 대해 강하게 반발하겠지만, 오늘날 그런 입장을 취하는 사람은 많지 않다. 카트라이트와 대립하는 물리주의자들도 인과력이라는 개념을 자연스럽게 사용하는 경우가 꽤 있다. 그렇다면 진정한 대립은 인과력이나 역량 개념의 도입 자체에 있는 것이 아니라 그 개념을 어떤 식으로 도입하느냐에 있을 것이다. 카트라이트는 물리적 존재자의 역량뿐만 아니라 사회적 역량 같은 온갖 다양한 역량들에 대해 이야기한다. 그녀는 이러한 다양한 역량

39) 우리의 언어에 그만큼 형상인과 목적인 개념이 자연스럽게 녹아들어 있다는 증거로도 볼 수 있다.

들이 실재적이라고 생각한다. 이는 그 자체로 특정한 존재론적 입장을 함축하는 것처럼 보인다.

그녀와 같은 입장에서는 거시적 대상이 그 대상임으로서 가지는 역량에 대해 부정할 아무런 이유가 없다. 그녀가 자주 드는 예로, 아스피린은 아스피린이라는 거시적 구조를 가지는 덕분에 두통을 완화하는 힘을 가진다고 말할 수 있다. 인간 행위자가 가지는 정신적 인과력에 대해서도 같은 식의 이야기를 할 수 있다. 그런데 그녀는 거시적 대상이 미시적 대상들로 이루어져 있다는 사실을 부정하지 않는다. 다시 말해서, 그녀는 비물질적 영혼 같은 존재를 도입하지 않는다. 사실 환원 불가능한 다양한 역량들의 실재성을 인정하기 시작하면 비물리적 존재를 도입하는 것은 별로 매력적인 선택지가 아닌데, 정신의 인과력을 설명하기 위해 비물리적 영혼을 도입하는 것까지는 허용될 수 있다고 쳐도, 다른 거시적 대상들이 가지는 거시적 인과력을 설명하기 위해 그에 해당하는 각각의 비물리적 존재자들까지 도입하는 것은 누구도 바라지 않을 그림이기 때문이다. 그래서 데카르트 본인의 경우에서 알 수 있듯이, 비물질적 영혼을 도입하는 사람은 보통 정신의 인과력 이외의 다른 거시적 인과력에 대해서는 환원주의적 태도를 취하는 경향이 있다.

어쨌든 거시적 대상들이 가지는 역량의 실재성을 인정하면서, 동시에 거시적 대상이 미시적 대상들로 이루어져 있다는 사실을 받아들인다면, 대상의 부분과 전체 사이의 관계를 묻는 물음에 대해 특정한 답을 줄 수밖에 없을 것이다. 왜냐하면 인과력과 존재의 지위 사이에는 불가분한 관계가 있다는 것이 자연스러운 생각이기 때문이다. 적어도 자연적 대상에 한하여, 거시적 대상이 거시적 대상으로서의 고유한 인과력을 가지고 있다고 말하면서 동시에 거시적 대상이 미시적 대상들의 단순한 집적체에 지나지 않는다고 말하는 것은 이상하게 들린다. 거시적 대상이 고유한 인과력을 가지고 있다면, 그것은 그것을 이루는 부분들의 단순한 합 이상의 것이어야 한다.

이와 같은 부분과 전체 사이의 관계에 대한 물음은 거의 철학사만큼이나 오래된 것이다. (아리스토텔레스도 『형이상학』, 7권, 1041b 등등에서

이 문제를 언급하고 있다.) 우리는 부분들로 이루어진 전체가 부분들을 그저 모아놓은 것인지, 아니면 그 이상의 어떤 것인지 묻는다. 모래 더미는 전자에 속하는 것 같고, 생명체는 후자에 속하는 것 같다. 이 두 가지 유형의 전체 중에서 전자는 전통적으로 흔히 단순한 집적체(mere aggregate)라고 불리었고, 후자는 통합된 전체(integrated whole) 혹은 유기적 통일성(organic unity)이라고 불리었다.⁴⁰⁾ 만약 환원 불가능한 거시적 역량을 진정으로 받아들인다면, 자연에 후자에 속하는 대상들이 실제로 존재한다는 점도 인정하는 것이다.

거시적 대상의 고유한 인과력을 인정하는 것과 그것이 대상의 구성에 대해 함축하는 바 사이의 연관성에 대해 카트라이트는 역량 개념을 도입할 당시에는 별로 의식하지 못했던 듯하다. 전통적인 형이상학적 문제들이 그녀의 주된 관심사가 아닌 탓이 클 것이다. 역량들의 실재론이 부분과 전체 사이의 관계에 대해 함축하는 바에 그녀가 그나마 의식하기 시작한 것은 비교적 최근의 일인 것으로 보인다. “부분들의 배열은 부분들 각각이 아니라 전체에 내재적이다. [법칙적-인용자] 기계의 배열은 부분들 각각에 의해 소유되지 않는 창발적 힘을 가질 수 있다(Cartwright and Pemberton 2013: 111)”고 간략히 언급함으로써 이 문제에 대한 견해를 살짝 내비치기는 하였지만, 깊이 다루고 있지는 않다. 하지만 조금 전까지 진행된 논의가 옳다면, 카트라이트의 세계관을 압축하는 ‘얼룩덜룩한 세계’라는 표현은 그녀가 의도한 것보다 더 많은 것을 의미한다. 이 표현은 일차적으로 반토대주의를 표방하기 위해 사용된다. 미시 물리학은 모든 지식의 토대를 이루지 못하며, 환원 불가능한 다양한 이론들, 인과적 설명 유형들과 더불어 “법칙들의 누더기(patchwork of laws)”를 이루는 한에서만 현실을 설명할 수 있다. 왜냐하면 다양한 역량들이 이미 현실에서 작용하고 있기 때

40) 최근에 쉐퍼는 근거 개념을 사용해 이 문제를 정식화한 바 있다(Schaffer 2009: 374). 그의 정식화에 따르면, 단순한 집적체에서는 전체가 부분에 근거하고, 통합된 전체에서는 부분이 전체에 근거한다. 여기에 한 가지 분류를 더 추가하자면, 인공품은 마치 통합된 전체인 것처럼 보이지만, 실제로는 단순한 집적체에 지나지 않는 독특한 대상이라고 할 수 있다. 물론 인공품의 각 부분들 자체는 통합된 전체인 자연적 대상들일 것이다. 인공품이 가지는 것처럼 보이는 고유한 인과력은 다른 자연적 대상, 예를 들자면, 제작자인 인간 등등이 가지는 본래적인 인과력에서 파생된 것으로 보아야 할 것이다.

문예, 그것들을 제어하기 위한 지식들 역시 필요하기 때문이다. 이런 그림이 형이상학 관점에서 흥미로운 이유는 다양한 역량들이 실재적이기 위해서는 역량들 사이의 배열과 질서가 문제가 된다는 점이다.⁴¹⁾ 이 문제는 실체에 관해 논하는 이 글의 4장에서 다시 중요하게 다루어질 것이다.

역량 개념의 도입이 오늘날 어떤 사람들에게 문제가 된다면, 그 이유는 그 개념 자체가 아니라 부분적 역량을 초월하는 새로운 전체의 역량을 도입한다는 데 있을 것이다. 이런 부류의 입장에 대한 대표적인 반론이 배제 문제인데, 지금까지 보았듯이 배제 문제를 이루는 전제들 중 하나인 물리 영역의 인과적 폐쇄성을 논파했으므로 받아들이는 데에 큰 문제는 없다.

3.4. 결론

정신적 원인이 물리적 결과를 일으킬 수 있다는 데 대해서는 거의 모든 사람들이 의심하지 않는다. 하지만 물리학자들은 많은 물리적 현상에 대해

41) 반 굴릭은 거시적인 물리적 대상들이 단순한 집적체 이상으로 보인다는 점이 곧 거시적 대상들이 진정한 인과력을 가지고 있다는 것을 믿을 만한 근거라고 여긴다는 점에서 반대 경로를 취하고 있다. 그에 따르면 대상의 거시적 구조 혹은 패턴은 다음과 같은 특성들을 보여준다. “(1) 그와 같은 패턴들은 세계의 회귀적이고 (recurrent) 안정적인 측면들이다. (2) 많은 그와 같은 패턴들은 그것들의 기저에 있는 물리적 구성 요소들의 변화에도 불구하고 안정적이다. 그 패턴들은 그것들의 구성 요소가 보존되지 않음에도 불구하고 보존된다(예: 허리케인이나 풀잎). (3) 많은 그와 같은 패턴들은 그것들을 퇴화시키거나 파괴할지도 모르는 교란하는 물리적 힘들에 직면하여 자기 유지적 혹은 자기 재생산적이다(예: DNA 패턴들). (4) 그와 같은 패턴들은 그것들의 구성 요소들이 가지는 어떤 인과력이 활성화되거나 활성화될 법한지에 영향을 줄 수 있다. 주어진 물리적 구성 요소들은 많은 인과력들을 가질지 모르지만, 주어진 상황에서 오직 그들 중 몇몇 부분 집합만이 활성화될 것이다. 그것이 부분인 더 큰 맥락(즉, 패턴)은 그것의 인과력들 중 어떤 것이 활성화될지에 영향을 미칠지 모른다. (예를 들어, 시약의 활성화는 그 시약과 더불어 화합물을 형성하는 촉매 효소의 현전에 의해 영향을 받는다.) 따라서 전체는 그것의 부분들의 어떤 단순한 기능도 아닌데, 왜냐하면 전체는 적어도 부분적으로라도 그것의 부분들에 의해 어떤 기여들이 만들어질지 결정하기 때문이다. (5) (4)와 같은 인과력의 선택적 활성화는 많은 경우 패턴 그 자체의 유지와 보수에 기여한다”(Van Gulick 1993: 251-52). 반 굴릭이 보기에는 미시적 입자들도 이러한 거시적 패턴들 가운데 하나에 불과하다는 점이 무엇보다 중요하다.

언제나 충분한 원인을 제시하는 것 같고, 정신적 원인은 인과력을 가지고 있어도 할 일이 없는 것처럼 보인다는 것이 배제 문제의 요점이었다. 이제 우리는 물리학자들이 많은 물리적 현상에 대해 언제나 충분한 원인을 제시하는 것처럼 보이는 이유가 모델 구성 이후만 보고 있었기 때문이라는 점을 안다. 모델 구성 이전까지 시야에 넣는다면, 물리 영역의 인과적 폐쇄성이 아니라 오히려 개방성이 경험적 지지를 받는 것으로 보일 것이다. 모델 구성 이전에 우리의 경험에 주어지는 세계는 온갖 다양한 인과력들이 온갖 다양한 결과를 일으키는 ‘얼룩덜룩한 세계’이다. 본질적으로 제한된 적용 범위를 가질 수밖에 없는 이론들은 각자의 목적과 관심에 따라 세계의 한 국면을 포착하여 설명하려고 시도한다. 따라서 이론 혹은 인과적 설명 유형들 역시 다수일 수밖에 없다.

물리 영역의 인과적 폐쇄성 원리가 경험적 지지를 얻지 못한다면, 굳이 무리하게 물리주의를 주장할 필요성도 사라진다. 그럼에도 불구하고 궁극적으로 세계는 신적 관점에서 보면 하나의 수준으로 축소될 수도 있지 않을까? 물론 그렇게 믿고 싶다면 믿을 수도 있다. 하지만 우리에게서는 행위자로서의 우리 자신이 가지고 있는 능력에 대한 자명하고도 근본적인 믿음과 잘 조화되기도 어렵고 별다른 경험적 지지도 받지 못하는 그러한 토대주의적 신념을 가질 좋은 이유가 없다.

4. 속성 이원론과 실체 이원론, 그리고 실체 다원론

이제까지의 논의들을 통해 우리는 다양한 인과력들의 실재성을 인정함으로써 부수현상론을 피하면서도 배제 문제를 일으키지 않을 수 있는 방법에 대해 알아보았다. 이제 정신적 속성이 정신적 속성으로서 가지는 인과력, 그리고 다른 거시적 대상들이 그 대상임으로서 가지는 인과력들은 확실한 존재론적 지위를 보장받게 되었다. 이는 물리주의를 거부하는 비주류 이론들이 공통적으로 해결해야 하는 문제를 해결했다는 것을 의미한다. 하지만 다른 관점에서 보자면, 물리주의를 거부하면서 시작하는 입장들 사이에서 어떤 입장이 가장 설득력 있는 입장인지 밝혀야 하는 과제가 아직 남아 있다는 뜻이기도 하다.

이번 장의 논의는 실체 이원론자와 속성 이원론자 사이의 논쟁을 살펴보는 것으로 시작된다. 여기서 속성 이원론은 물리적 실체 일원론과 결합된 입장을 의미한다.⁴²⁾ 실체 이원론자는 우리가 일단 속성 이원론을 받아들이면, 실체에 관한 어떤 이론을 가지고 있든지 실체 이원론까지 받아들일 수밖에 없다고 주장한다. 반면에 속성 이원론자는 우리가 실체에 대해서는 일원론을 고수할 수 있다고 주장하면서 실체 이원론자를 논파하려 한다. 나는 실체 이원론이 제기하는 근본적인 물음에 속성 이원론자가 제대로 답하지 못했다는 점을 보여줄 것이다.

우리가 실체 일원론을 거부한다 하더라도 그러면 어떤 식으로 실체가 여럿일 수 있는가에 관한 논의는 아직 열려 있다. 데카르트식 순수 실체 이론에 의해 실체의 유형을 늘릴 수도 있고, 다른 식으로 늘릴 수도 있다. 나는 데카르트식 실체 이론에 대해 제기되는 대표적인 문제들이 결정적이

42) 비환원적 물리주의가 속성 이원론으로 분류되는 경우도 있으며, 이번 장에서 속성 이원론에 가해지는 비판이 비환원적 물리주의에도 적용된다고 볼 수도 있다. 하지만 이번 장에서 내가 염두에 두고 있는 속성 이원론은 2장에서 살펴본 베넷과 같은 부류의 비환원적 물리주의가 아니라 창발적 속성 이원론과 같은 입장들이다.

라고까지는 생각하지 않지만, 훨씬 더 매력적인 다른 대안이 있다는 점을 보여주하고자 한다. 특히 아리스토텔레스주의적 실체 이론은 이제까지 제기된 문제들을 말끔하게 해결하면서 실체 다원론을 설득력 있게 제시하도록 해 줄 것이다. 이 글에서 전개된 아리스토텔레스주의적 실체 이론은 창발론의 일종으로 해석될 수 있다. 창발론은 종종 수반을 필수 요소로 가지는 것처럼 이해되기도 하였지만, 관련된 최근의 논의들을 통해 드러난 바는 반드시 그러할 필요는 없다는 것이다.

4.1. 속성 이원론에서 실체 이원론으로

심신 문제에 관한 최근의 논의들에서 물리주의가 지배적인 경향이라는 것은 주지의 사실이다. 놀라운 점은 물리주의를 거부하는 많은 논자들까지도 실체 차원에서는 오로지 물리적 실체만이 존재한다는 점을 당연하게 받아들인다는 것이다. 일반적으로 ‘속성 이원론’이라 불리는 입장이 바로 그러한 부류에 속한다. 그래서 일반적인 속성 이원론은 물리적 실체 일원론이기도 하다. 속성 이원론은 정신적 속성이 환원 불가능할 뿐 아니라 존재론적으로도 근본적인 지위를 가지고 있다고 본다는 점에서 전형적인 물리주의와 차별화된다. 하지만 속성 이원론은 어쨌든 실체는 물리적 실체 한 유형뿐이라고 주장한다는 점에서는 물리주의와 같은 노선을 따르고 있다. 마치 가능한 실체의 유형에 관해서는 합의가 이미 이루어진 것처럼 말이다.

조금만 생각해 보면, 이는 이상한 일이라는 것을 알 수 있다. 인간은 정신적 속성과 물리적 속성을 모두 가지고 있다. 그런데 왜 인간은 물리적 실체인가? 정신적 속성은 어째서 실체의 유형을 결정하는 데 참여하지 못하는가? 정신적 속성, 조금 더 자세히 말해서 이러저러한 특정한 정신적 속성이 아니라 정신적 속성 일반은 인간이 우연히 일시적으로 얻거나 잃을 수 있는 속성인 것처럼 보이지 않는다. 그것은 인간이 인간인 한에서 반드시 가지고 있어야 하는 본질적 속성인 것처럼 보인다. 게다가 정신적

속성은 인간을 다른 존재들로부터 구분시켜주는 속성이기도 하다. 정신적 속성이 본질적이고, 존재론적으로도 근본적이라는 점을 받아들이면서도 속성 이원론자는 왜 인간이 물리적 실체라고 주장하는가? 이는 속성 이원론자들이 잘 묻지 않는 아주 단순하고 기본적인 물음이다. 이 문제에 대해 속성 이원론자의 입장에서 그럴듯한 답변을 내놓을 수 있을까?

최근에 슈나이더(Schneider 2010)는 위의 물음을 집요하게 추적하는 논문을 발표했다. 이 논문에서 그녀는 속성 이원론자들이 실체에 관한 어떤 이론을 가지고 있는지 속성 이원론과 물리적 실체 일원론을 결합시키는 것이 실패한다는 것을 보여주고자 한다. 그녀의 논증은 다음과 같이 요약될 수 있다.

우리가 실체에 관해 다발 이론을 가지고 있다고 해 보자. 그러면 실체는 어떤 속성들(보편자이든 트롭이든)의 다발에 지나지 않는다. 물론 다발 이론에서도 함께 묶여 있는 모든 속성이 실체를 유형화하지는 않는다. 하지만 정신적 속성은 우연적 속성이 아니다. 따라서 인간이라는 실체는 정신적 속성과 물리적 속성의 다발일 텐데, 어째서 실체를 물리적 실체라고 단정할 수 있는가? 속성 이원론자는 이 물음에 선결 문제 요구의 오류를 범하지 않고 답할 수 있어야 한다.

이번에는 우리가 실체에 관해 기체(substratum) 이론을 가지고 있다고 해 보자. 다발 이론의 경우와는 달리, 이 경우 속성 이원론자들은 기체 자체가 물리적인 것이라고 주장하고 싶어한다. 그런데 기체란 무엇인가? 이 개념에 대한 로크의 유명한 비판("I know not what")에서도 알 수 있듯이, 기체는 종종 신비한 것으로 여겨진다. 그럼에도 불구하고 기체를 이해하고자 한다면, 두 가지 해석이 가능하다. 강한 개성 원리(Haecceitism): 서로 다른 두 기체는 속성과는 다른 독특한 내적 본질을 가지는 덕분에 서로 개별화된다. 2) 약한 개성 원리: 기체들은 내적 본질을 가지지 않으며, 오로지 수적으로만 구별된다.

약한 해석에 따를 경우, 실체는 다발 이론의 경우와 마찬가지로 그것이가지는 속성에 의해 유형화된다고 보아야 한다. 그렇지 않으면 어떤 기체를 물리적 기체라고도 할 수 없을 것이다. 하지만 실체가 속성에 의해 유

형화될 경우, 최초로 제기한 문제는 여전히 남아 있게 된다. 강한 해석에 따르더라도 이 기체를 물리적 기체라고 말하기는 어렵다. 왜냐하면 그것이 가지고 있는 독특한 내적 본질은 속성과 다르다는 것이 전제되어 있기 때문이다. 그것을 물리적 기체라고 한다면, 그것은 그것의 속성에 의해 개별화된다는 뜻인데, 이는 강한 개성 원리와 부합하지 않는다.

결국 속성 이원론자는 물리적 실체 일원론을 옹호하는 데 실패한다. 그리고 만약 속성 이원론자가 혼합형 실체 개념, 즉 정신적 속성과 물리적 속성을 모두 본질적으로 가지고 있기 때문에 순수하게 물리적이라고는 할 수 없는 실체 개념을 도입한다면, 그는 실체 이원론자가 되는 것이다. 왜냐하면 그에게는 최소한 혼합형 실체와 순수 물리적 실체라는 두 가지 종류의 실체가 있기 때문이다.

이상이 슈나이더가 물리적 실체 일원론을 비판하는 논증의 핵심이다. 그리고 당연한 말이지만 이 논증은 심신 문제를 넘어서서 일반화될 수 있다. 한 가지 흥미로운 점은 그녀의 논증이 실제 속성 이원론자의 주장을 비판하는 형식으로 되어 있지 않다는 점이다. 그녀의 논증은 오로지 가상의 속성 이원론자가 취할 수 있는 입장들을 예상하고 반박하는 형태로 되어 있다. 그만큼 물리적 실체 일원론이 아무런 문제 의식 없이 당연한 것으로 가정되어 있는 현실을 반영하고 있는 것이라 할 수 있다.

다발 이론과 기체 이론이 실체에 관한 이론의 전부는 아니다. 슈나이더 자신도 아리스토텔레스주의적인 실체 개념을 고려하고 있다(Schneider 2010: 6, n7). 하지만 아리스토텔레스주의적인 실체 개념은 물리적 실체 일원론을 주장하기 위한 용도로 적합하지 않다. 이에 대해서는 3, 4절에서 좀 더 자세히 논하도록 하겠다.

4.2. 반론과 응답

최근에 양(Yang 2015)은 슈나이더가 제기한 문제에 대해 속성 이원론자의 입장에서 답하고자 시도한 바 있다. 그의 반론은 크게 두 가지로 나

놓 수 있다. 하나는 슈나이더가 미처 고려하지 않은 실체에 관한 이론까지 고려하면, 속성 이원론이 문제를 해결할 수 있다는 것이고, 다른 하나는 슈나이더가 고려하고 있는 실체에 관한 이론들 내에서도 문제없이 속성 이원론을 주장할 수 있는 길이 있다는 것이다.

우선 첫 번째 반론을 살펴보자. 양은 슈나이더가 고려하고 있는 실체에 관한 이론들이 모두 구성요소 존재론으로 분류할 수 있다고 하면서, 관계적 존재론에서는 그녀가 제기하고 있는 것과 같은 문제가 발생하지 않는다고 주장한다(Yang 2015: 4-5). 여기서 그가 대비시키고 있는 구성요소 존재론과 관계적 존재론이라는 것은 각각 무엇을 의미하는가?

아리스토텔레스에 의해 고취된 존재론들이 ‘구성요소’ 존재론인 반면, 플라톤에 의해 고취된 것들은 ‘관계적’ 존재론이라는 혼한 가정-아리스토텔레스의 형이상학이 구별되게 ‘내재적’인 반면 플라톤은 구별되게 ‘초월적’이라는 개념에 근거한 가정이 있다(…) [그리고] 내재주의자는 구체적인 대상들의 속성들을 바로 그 대상들의 성분인 것으로 보는 반면 초월주의자는 그것들을 대상들이 그것들과 예화라는 특별한 관계를 맺는 분리된 존재자들인 것으로 본다(Lowe 2012: 229-30, Yang 2015: 4에서 재인용).⁴³⁾

인용문에서 알 수 있듯이, 구성요소 존재론과 관계적 존재론은 속성과 실체 사이의 관계에 관한 대비되는 이론들이다. 여기서 실체가 속성들을 구성요소로 가진다는 것이 무슨 의미인지, 예화라는 특별한 관계를 어떻게 이해해야 좋을지 등등의 의문이 제기될 수도 있지만, 그것은 지금 이 논의에서는 중요하지 않다. 양도 그런 것들은 전혀 고려하고 있지 않다. 어쨌든 그는 관계적 존재론을 채택하면 속성 이원론이 안전하게 물리적 실체 일원론을 고수할 수 있다고 주장하는데, 그 이유가 문제적이다.

그에 따르면, 물리적 실체가 물리적 실체인 이유는 그것이 오직 물리적 대상들만을 부분으로서 갖기 때문이다. 물리적 실체는 물리적 속성이든 정

43) 로워(Lowe 2012)의 논문은 그 제목이 말해주듯이, 아리스토텔레스가 구성요소 존재론도 관계적 존재론도 아닌 제3의 존재론을 채택하고 있다는 것을 주장하고자 하고 있다.

신적 속성이든 모두 예화할 수 있지만, 관계적 존재론을 받아들인다면 그것들은 실체의 구성요소, 즉 부분이 아니기 때문에 문제를 일으키지 않는다(Yang 2015: 5-6).

내가 보기에 그는 다음과 같은 것을 주장하고 있는 것 같다. 우선 정신적 속성과 물리적 속성을 모두 예화하는 물리적 실체가 있다. 편의상 그 실체를 두뇌라고 해 보자. 그런데 두뇌는 배타적으로 물리적 속성만을 예화하는 물리적 부분들만으로 이루어져 있다. 따라서 두뇌는 물리적 실체이다.

그가 정말로 위와 같은 주장을 하고 있는 것이라면, 다음과 같은 문제가 제기될 것이다. 물리적 부분들이 두뇌를 이루고 있다고 할 때, 그 이루는 관계라는 것이 도대체 무엇인가? 그가 쓴 구절들을 읽어보면 두뇌가 마치 물리적 부분들의 단순 집적체인 것처럼 그려지는데, 어떻게 물리적 부분들의 단순한 집적에 불과한 두뇌가 환원 불가능하고 존재론적으로 근본적인 정신적 속성을 예화할 수 있는가? 그리고 정말로 두뇌가 물리적 부분들의 단순한 집적에 불과하다면, 두뇌를 물리적 실체라고 말하는 것은 엄격한 의미에서는 옳바르지 않다. 정말로 그러하다면 진정한 의미에서의 물리적 실체는 그 물리적 부분들이라 해야 할 것이다. 물리주의자들에게 진정한 실체는 미시 물리적 존재자들뿐인 것처럼 말이다. 당연한 말이지만 속성 이원론자는 이런 식의 입장을 취할 수 없다. 속성 이원론자의 입장에서 정신적 속성을 예화하지 않는 미시 물리적 존재자들만이 진정한 실체라면, 실체에 관한 기술이 완전히 주어져더라도 세계의 한 근본적인 측면인 정신적 속성에 관한 기술은 여전히 결여되어 있다는 뜻인데, 이는 실체라는 개념에 대해 우리가 가지고 있는 일반적인 이해에 비추어 볼 때 말이 되지 않는다. 실체에 관한 완전한 기술이 주어진다면, 세계의 근본적인 측면들에 관한 기술이 주어진다고 여기는 것이 일반적이다. 따라서 속성 이원론이 그 명칭에 걸맞는 입장이 되려면 진정한 의미에서의 실체가 물리적 속성과 정신적 속성을 모두 예화하고 있다고 주장해야 할 것이다. 그러면 다시 슈나이더가 처음에 제기한 문제가 되돌아온다.

양의 주장은 다르게 해석될 수도 있다. “실체를 물질적 대상으로 만드는

것은 오직 물질적 대상들만을, 특히 상식과 경험 과학에 의해 부여된 것들만을 (적절한 혹은 부적절한) 부분으로서 가지는 것이다. 하지만 물질적 실체는 물리적 속성과 정신적 속성 모두를 예화할 수 있고, 그것들 중 어느 것도 어떤 의미에서도 그 대상의 구성요소가 아니다”(Yang 2015: 5-6)라는 말은 마치 물질적 대상들 가운데 하나인 물리적 실체가 물리적 속성과 정신적 속성의 예화에 앞서 주어져 있을 수 있다고 말하는 것처럼 들리기도 한다. 어떻게 대상이 속성 예화 이전에 ‘물질적’이라는 규정을 가질 수 있는가? 그는 상식과 경험 과학이 그러한 일을 하고 있다고 말하는 것 같다. 하지만 우리가 관행적으로 어떤 대상들을 물질적이라고 부른다고 해서 그것이 지금 이 논의에서 어떤 대상을 속성 예화 이전에 물질적인 것으로 규정하게 해 준다고 말하는 것은 논의의 차원을 혼동하는 것에 불과하다. 속성 예화 이론을 진지하게 받아들인다면, 속성 예화 이전의 대상들은 그것 자체로는 순수한 무규정적 대상이라고 보아야 한다(Lowe 2012: 234).⁴⁴⁾

이상의 논의에서 알 수 있듯이, 구성요소 존재론과 관계적 존재론을 구분하는 것은 속성 이원론자의 입장에서 문제를 해결하는 데 별로 도움이 되지 않는다. 물론 앞서 말했듯이, 양에게는 한 가지 반론이 더 남아 있는데, 이는 실체 이론과 독립적으로 전개되는 반론이다.

우선 그는 개인의 자기동일성이 그의 생물학적 기관들에 의해 규정된다는 동물주의(animalism)를 받아들인다. 그는 동물주의를 물리주의의 일종으로 여기고 있다. 즉, 그에 따르면 정신적 속성은 인간의 본질적인 속성이 아니며, 지속 조건에도 관여하지 않는다. 이런 입장에 따르면, 가령 식물인간은 어떠한 정신적 속성도 가지고 있지 않고, 가질 가능성조차 박탈당했다고 생각되지만, 자기동일성을 유지하고 있는 것이다. 결국 동물주의가 옳다면, 정신적 속성은 근본적이기는 하지만 본질적 속성은 아니므로 실체를 유형화하는 데 개입하지 않는다는 것이다.

동물주의는 논란의 여지가 너무 많은 입장이다. 정신적 속성이 인간의 본질적 속성이라는 생각은 쉽게 거부하기 어려운 직관을 동반하고 있다.

44) 순수한 무규정적 대상 혹은 순수한 질료는 물리주의자들이 말하는 물리적 대상이 아니다.

정말로 동물주의가 옳다면, 개인의 자기동일성은 태아일 때부터 식물인간 상태가 될 때까지 변함이 없다는 것인데, 그렇다면 왜 우리가 낙태나 안락사 문제로 그토록 격렬하게 논쟁하는지도 해명되어야 할 것이다.

게다가 설령 동물주의가 맞다고 하더라도 물리적 실체 일원론을 주장하기까지는 아직 갈 길이 멀다. 양은 동물주의가 아무런 문제 없이 물리주의의 일종으로 간주될 수 있다고 생각하는데, 이제까지 진행된 이 글의 논의를 되새겨 보면, 그것은 별로 그럴듯하지 않은 가정이다. 동물적 속성이 물리적 속성으로 환원되는가? 오히려 동물적 속성은 환원 불가능하게 인과적 설명에서 나타나는 것처럼 보인다. 부수현상론을 피하고 싶다면, 동물적 속성의 존재론적 지위를 인정해야 할 것이다. 또, 동물적 속성은 동물의 본질적 속성이라는 점은 의심의 여지가 없다. 그렇다면 왜 동물적 속성과 물리적 속성 양자를 본질적으로 가지는 동물은 물리적 실체인가? 결국 동물주의에 호소하여 속성 이원론을 옹호하려는 시도는 그저 문제를 다른 곳으로 옮겨놓았을 뿐이다.

4.3. 순수 실체 이론 대 혼합 실체 이론

이로써 물리적 실체 일원론과 속성 이원론이 결합되기 어려운 입장이라는 점이 분명해졌다. 이제 우리는 실체 이원론을 진지하게 다루어야 할 때가 되었다. 종종 간과되는 사실이기는 하지만, 여러 가지 이원론들이 있을 수 있다.

우선 실체 이원론의 대표 격인 데카르트식 이원론이 있다. 데카르트식 이원론은 말하자면 순수 실체 이원론이다. 이 이론에서 정신적 실체는 어떠한 물리적 속성도 가질 수 없으며, 물리적 실체는 어떠한 정신적 속성도 가질 수 없다.

이와 같은 순수 실체 이원론은 가능한 이원론들 중에서도 가장 극단적인 입장이라 할 수 있다. 그렇기 때문에 그만큼 잘 알려진 문제점들을 가지고 있다. 이미 데카르트 당대에 제기된 실체들 간 상호작용 문제가 있

고, 그 문제의 현대적 판본인 짝지음(pairing) 문제가 있다. 이런 문제들이 순수 실체 이원론을 완전히 논파할 정도로 강력한 것이냐에 대해서는 논란의 여지가 있다. 실체들 간 상호작용이든 특정한 정신적 실체와 물리적 실체의 짝지음에 대해서든 그 관계가 원초적이라고 말함으로써 더 이상의 설명을 거부하는 것도 순수 실체 이원론자의 입장에서 제시될 수 있는 한 가지 대답이기 때문이다.

그렇지만 순수 실체 이원론과 그것이 도입하는 비물질적 영혼의 개념은 여러 가지 면에서 오늘날 많은 사람들이 보기에 별로 매력적이지 않다는 점은 인정되어야 할 것이다. 무엇보다 정신적인 것이 육체에 의존하는 것으로 보이는 많은 현상들이 있다. 감각 기관을 통한 인식도 그 중의 하나이다. 감각 기관을 통한 인식 없이 정신이 어떤 사고를 할 수 있을까? 이런 물음들에 대해서도 순수 실체 이원론의 입장에서 주어질 수 있는 대답들이 있겠지만, 사람들에게 얼마나 그럴듯한 것으로 받아들여질지는 의문이다. 그리고 어쨌든 순수 실체 이원론자들도 비물질적 영혼이 시공간상의 특정한 위치에 있는 한 육체와 긴밀하게 연결되어 있다는 점은 받아들인다. 그렇다면 물리적 속성을 완전히 결여하고 있는 실체의 개념을 따로 도입하기보다는 정신 자체가 모종의 물리적 속성을 가지고 있다고 하는 편이 더 간단하지 않을까? 이 점은 특히 우리가 고려하고 있는 것이 단지 정신적 속성만이 아니라는 점을 상기하면 더욱 그럴듯하게 느껴진다. 데카르트에게 본질적인 속성은 오직 두 가지 뿐이었다. 하지만 우리는 그렇게 생각할 수 없다. 지난 장들에서 보았듯이, 우리는 연장적 속성 혹은 그 현대적 대응물인 물리적 속성으로 환원되지 않는 다양한 본질적 속성들을 알고 있다. 이 다양한 본질적 속성들에 대응하는 각각의 비물리적 실체들이 있다고 해야 할 것인가? 영혼불멸 같은 특별한 종교적 교리와의 친화성을 고려하지 않는다면, 순수 실체 이론은 다른 이원론들에 비해 매력적인 대안은 아니다.⁴⁵⁾

우리가 순수 실체 이론을 포기한다면 다른 이원론적 대안들이 눈에 들어오기 시작한다. 슈나이더는 한 예시로서 다발 이론과 연계된 혼합형 실

45) 정신적 실체가 어떠한 물리적 속성도 가질 수 없다는 점을 보이려는 데카르트의 논증은 오늘날 받아들여지기가 힘들다(Lowe 2006: 7-8).

체 개념을 제시한다(Schneider 2010: 14). 혼합형 실체 개념에 따르면, 정신적 실체 혹은 비물리적 실체는 물리적 속성 또한 가질 수 있다. 이 혼합형 실체를 정신이라고 한다면, 정신은 특정한 공간상의 위치를 가지는데, 그 위치는 복잡한 물리적 실체(예: 육체)와 일치하지만, 물리적 실체와는 다른 지속 조건을 가진다는 점에서 구별되는 실체이다. 물질 대사 등에 의해 육체의 물리적 구성이 변하더라도 정신의 자기동일성은 유지된다는 점은 분명하다. 그렇지만 여기서 정신과 육체 사이에는 법칙적 필연성이라는 의존 관계는 있는 것으로 설정되어 있다. 즉, 육체의 존재는 정신의 존재를 위해 법칙적으로 충분하다.⁴⁶⁾

슈나이더가 한 예시로서 제시하는 입장이기 때문에 그녀 자신이 얼마만큼 이러한 입장을 믿고 있는지는 미지수이지만, 이러한 식의 입장에 대해 일반적으로 지적할 사항이 몇 가지가 있다. 우선 혼합형 실체를 도입한다면, 순수 실체 이론에 제기되는 많은 문제들이 제기되지 않으리라는 점은 분명하다. 그 점만으로도 순수 실체 이론에 비해 혼합형 실체 이론을 선택할 동기가 될 수도 있을 것이다. 그런데 슈나이더는 다발 이론과 연계된 혼합형 실체 이론을 선택할 경우, 어째서 이질적인 정신적 속성과 물리적 속성이 단일한 다발에 속할 수 있는지를 설명할 수 있어야 한다고 한 바 있다(Schneider 2010: 8). 그런데 정작 그녀 자신은 이 물음에 대한 답을 주지 않고 있다. 그것이 정말로 문제가 되는 물음이라면, 하다못해 두 속성이 한 다발로 묶일 수 있는 것은 원초적 사실이라는 식의 답변이라도 있어야 할 것이다.

46) 슈나이더는 이런 입장을 일종의 창발적 실체 이원론으로 분류하고 있다. 본래 그녀가 든 예시에서 정신적 실체와 짝을 이루는 복잡한 물리적 실체는 두뇌이다. 이는 심각한 검토 없이 근래의 유행에 따른 것으로 보이며, 그녀가 취하고 있는 입장에 본질적인 것은 아니라고 여겨지므로 중립적인 용어인 ‘육체’로 대체했다. 비교적 최근에 제기된 체화된 인지 관련 논의가 참이라면, 두뇌는 정신을 위한 법칙적 필요조건이기는 하지만 충분조건은 되지 못할 것이다. 그리고 육체조차도 환경을 필요로 하기 때문에 역시 충분조건은 되지 못할 것이다. 물리적인 것에 대한 정신적인 것의 의존 관계는 명백하지만, 그 관계는 두뇌와 정신 사이의 법칙적 필연성 관계보다는 복잡한 것이어야 할 것이다. 정신이 육체에 의존적이라는 점을 말하고자 하기 위해 그것이 두뇌와 같은 국지적인 시공간상의 위치를 점유한다고까지 말할 필요는 없다. 속성 이원론을 넘어서서 실체 이원론을 받아들이려는 입장에서는 더욱 그러하다.

또, 정신적 실체와 물리적 실체가 같은 시공간상의 위치를 점하고 있는 독립적인 실체들이라고 말하는 것은 부자연스럽게 들린다. 차라리 정신적 속성과 물리적 속성(혹은 비정신적 속성)을 가지고 있는 인간이라는 실체가 하나 있을 뿐이고, 비정신적 속성만을 가지고 있다고 간주되는 실체는 그것이 무엇이든 일차적 실체인 인간으로부터의 추상물이라고 하는 편이 자연스럽게 들린다. 이렇게 말하는 것이 유익한 이유가 있다. 인간을 구성하는 물리적 부분들(예: 분자들)은 그것이 인간을 구성한다는 맥락 바깥에서 전형적으로 보여주는 것과는 다른 성질을 보여준다. 카트라이트의 법칙적 기계라는 개념을 이용해 말하자면, 인간은 일반적으로 그것의 물리적 부분들이 가지는 역량들의 전형적인 표현을 관찰하기에 적합한 법칙적 기계가 아니다. 오히려 인간이라는 법칙적 기계 속에서 그것의 물리적 부분들이 가지는 역량들은 정신이 가지는 역량을 표현하도록 배열되고 조정된다. 예컨대, 내 몸을 이루는 분자들의 움직임은 내가 가진 특정한 동물적 기능, 나아가 내가 의욕하는 바를 고려하지 않으면 이해 불가능하다. 실체의 중요한 특징은 독립성이므로, 전체에 의존적인 부분들은 실체가 아니거나 적어도 덜 실체적이라고 해야 할 것이다. 물론 실체 역시 그것의 구성요소들에 의존하기는 하나, 그 의존성은 부분들이 전체로서의 실체에 의존하는 것과는 비대칭적이다. 다시 법칙적 기계 개념을 이용해 말하자면, 인간이라는 법칙적 기계 속에서 각 부분들은 그것들이 가진 역량의 전형적인 표현을 제약당하는 방식으로 전체에 의존하지만, 전체는 그것이 가진 역량을 전형적으로 표현하도록 부분들에 의존한다.⁴⁷⁾ 이는 다르게 보자면, 인간을 이루는 물리적 부분들은 그것이 가진 역량을 전형적으로 표현할 수 있는 법칙적 기계 속에서는 일차적 실체라고 볼 수 있다는 뜻이기도 하다.⁴⁸⁾

47) 로빈슨(Robinson 2014: 2.2.2)은 아리스토텔레스주의자의 입장에서 이를 다음과 같이 표현한다. 질료(추상적 사물이 아니라 실체가 무엇이나에 따라 그것을 구성하는 부분들)는 질료화된 사물이 무엇을 할 수 있는지를 제약하는 반면, 실체적 형상은 그것의 질료가 정확히 어떻게 행동하는지를 규정한다.

48) 루는 이를 잠재적 실체라는 개념을 통해 설명한다. “물리적 단순체들은 한 유기적 시스템과 독립해서도 존재 가능하다. 그런 의미에서 이것들은 진정한 실체들일 수 있다. 그러나 다른 한편으로 물리적 단순체들은 살아 있는 생명체 안에 있다는 의미에서 하위 등급의 존재론적 위상을 가진다(…) 살아 있는 생명체 안에 있을 때 물

위에서 제기한 문제들은 슈나이더가 실체 이론으로서 다발 이론을 택하기 때문에 생겨나는 것들이다. 일반적인 다발 이론은 이 문제들에 대한 좋은 답을 주기가 어렵다. 아마도 그녀는 신비한 기체를 도입하는 것보다는 나은 선택이라고 생각해서 다발 이론을 채택했을 것이다. 이 대목에서 제기되는 한 가지 의문은 그녀가 아리스토텔레스주의적 실체 이론을 알고 있음에도 불구하고 어째서 그 가능성을 진지하게 탐구하지 않았느냐는 점이다(Schneider 2010: 6, n.7). 만약 우리가 다발 이론이나 기체 이론이 아닌 아리스토텔레스주의적 실체 이론을 채택한다면 데카르트식 순수 실체 이론에 대한 더 나은 대안을 제시할 수 있을 것으로 보인다.

4.4. 아리스토텔레스주의적 실체 이론과 창발론

아리스토텔레스주의에서 정신은 육체의 형상이라고 흔히 말해진다.⁴⁹⁾ 형상은 간단히 말하자면 질료를 (적어도 자연적 대상에 한해서) “유기화(조직화)하는 원리(organizing principle)”인데, 우리는 이미 이 글에서 이에 해당하는 내용들을 다룬 바 있다. 살아 있는 인간은 무언가로 이루어져 있다. 여기서 그 ‘무언가’에 해당하는 것이 질료인데, 편의상 물리적 존재자들과 해 보자. 이 물리적 존재자들이 결합되어 ‘살아 있는’ 인간이 되려면, 아무렇게 결합되어서는 안 되고 특정한 방식으로 결합되어 있어야 할 것이다. 즉, 살아 있는 인간을 이루는 물리적 존재자들은 단순한 집적체가 아니라 통합된 전체가 되도록 결합되어야 한다. 이렇게 부분들이 통합된 전체를 이루도록 결합되는 방식 자체가 곧 형상이다.⁵⁰⁾ 우리는 보통 이와 같은 식으로 통합된 전체를 진정한 하나의 대상으로 취급

리적 단순체들은 단지 잠재적이거나virtual 가능한potential 실체들이다”(루 2010: 245).

49) 아리스토텔레스 해석가들 사이에서 일차 실체를 형상으로 보아야 하는지 아니면 질료형상 복합체인 개별자로 보아야 하는지에 대한 논의가 있는데, 이 글에서는 편의상 개별자를 일차 실체로 보도록 하겠다.

50) 기체 이론은 이 독특한 결합 방식 자체를 마치 별도의 대상인 것처럼 물화하여 신비한 존재자로 만든 결과라고 볼 수 있다.

하므로, 형상은 자연스럽게 어떤 대상이 무엇인지를 결정하는 역할도 한다.

다발 이론을 채택한 혼합형 실체 이론에서는 어째서 정신적 속성과 물리적 속성이 한 다발에 속할 수 있는지에 대한 좋은 설명을 주지 못한다는 것이 문제가 된 바 있다. 일반적인 다발 이론에서는 속성들이 우선적으로 주어지며, 실체는 환원적 범주에 불과하다. 따라서 다발 이론에서는 정신적 속성과 물리적 속성이 주어졌을 때 실체는 그저 그것들이 하나로 묶여 있다는 것만을 말할 수 있을 뿐, 왜 그런지에 대한 더 이상의 설명은 줄 수 없다. 반면에 아리스토텔레스주의적 실체 이론에서 정신적 속성(개별적인 믿음이나 욕구 같은 것이 아니라 정신적 속성 유형)은 물리적 속성과 같은 수준에서 하나로 묶이기를 기다리는 속성이 아니라 부분들이 가지고 있는 물리적 속성들을 유기화하는 특별한 방식인 상위 속성이다.

정신이 물리적 속성들을 단순히 묶는 것이 아니라 ‘유기화’한다는 점은 다발 이론과 연계된 혼합형 실체 이론에 제기된 다른 문제를 해결하는 데 중요하다. 우리는 정신적 속성을 가지고 있는 인간이라는 실체와 정신적 속성을 결여하고 있는 것으로 생각된 물리적 실체인 육체가 서로 같은 수준의 실체가 아니라는 직관을 가지고 있다. 일반적인 다발 이론에서는 이 직관적인 구분을 포착할 방법이 없었다. 하지만 아리스토텔레스주의적 관점에서 보면, 정신적 속성을 결여하고 있는 육체란 적절한 방식으로 결합되지 않은 부분들인 단순한 집적체에 불과하다는 점이 금세 드러난다.⁵¹⁾

슈나이더는 자신의 입장을 일종의 창발적 실체 이원론으로 분류하고 있는데, 아리스토텔레스주의적 실체 이론 역시 통합된 전체가 가지는 존재론적으로 새로운 속성에 대해 말한다는 점에서 창발론의 일종으로 볼 수 있다. 그런데 창발이라는 개념을 정확히 어떻게 이해할 것인지는 현재까지도 논의가 진행 중인 까다로운 문제이다. 이 글에서는 한 가지 중요한 쟁점만 짚고 넘어가고자 한다. 창발에는 모종의 수반 관계가 필수적인 요소인 것처럼 이해되는 경우가 많다. 하지만 현재 창발 개념을 이해하는 두 가지 대표적인 방식을 잘 들여다보면 반드시 그래야 할 필요는 없다는 사실을

51) 스템프(Stump 2013: 67, n.67)에 따르면, 아퀴나스는 형상화되지 않은 ‘육체’는 형상화된 인간의 육체와 동음이의적 관계에 있다고 생각했다고 한다.

알게 된다. 고전적인 창발론에서는 수반 관계가 필수적인 요소로 취급되는 반면, 대안적인 창발론에서는 그렇지 않은데, 대안을 주장하는 논자들은 고전적인 입장이 어떤 이유에서 그렇게 생각하는지 그리고 그 대가가 무엇인지를 잘 드러내 보여준다.

브로드 같은 20세기 초 영국 창발론자들의 영향으로 고전적인 창발론은 흔히 공시적인 수반 개념을 통해 정식화되곤 했다. 창발 속성과 기저 속성에 수반하는데, 단 두 수준의 속성을 잇는 창발 법칙이 고정되는 한에서만 수반한다. 창발 법칙은 물리 법칙을 비롯한 다른 법칙들에 의해 형이상학적으로 필연화되지 않는다는 의미에서 근본 법칙이다. 간단히 말하자면, 기저 속성은 창발 속성을 위해 법칙적으로 충분하다.⁵²⁾ 웡은 이런 종류의 창발론을 “수반 창발론”(Wong 2006: 347)이라는 명칭으로 부르면서, 여기서 말하는 수반이 ‘순수한’ 수반이라는 점을 지적하고 있다. 이때 순수하다는 것은 인과적인 요소가 포함되어 있지 않다는 것을 의미한다.

최근에 창발이라는 개념에 대해 위와 같은 것과는 다른 이해를 제공하고자 하는 논자들은 고전적인 수반 창발론에 대해 의문을 제기하였다. 그들이 문제삼는 것은 그러한 수반 관계의 설명 불가능한 고유한(*Sui generis*) 성격이다. 고전적 창발론에서 존재론적으로 구별되는 두 속성 사이의 수반 관계는 수반 관계를 설명하는 일반적인 관계들인 구성, 동일성, 실현, 인과 등등의 어떠한 관계로도 설명 가능하지 않으며, 오로지 공시적이고 비인과적인 공변 관계를 말하는 근본적인 창발 법칙에만 근거한다(O'Connor and Wong 2005: 662, Wong 2010: 20). 이러한 이론에서 창발적 속성에 의한 하향식 인과는 오직 창발 법칙이 규정하는 패턴을 벗어나지 않는 한에서만 일어날 수 있다. “그것은 확실히 정합적이고 가능하지만, 고유한 형이상학적 결정 관계라는 영원한 신비와 깊은 설명적 보수주의라는 매수자 위험부담(*caveat emptor*)의 대가를 지불하고서만 그러하다”(Wong 2010: 20).

고전적인 창발론에서 수반 관계가 설명 불가능할 뿐만 아니라, 동시에 창발적 속성의 인과적 역할을 경직되게 만든다는 문제 의식 때문에 최근

52) 슈나이더가 염두에 두고 있는 것도 이러한 형태의 고전적인 창발론인 것 같다.

에는 “인과적 창발론”(Wong 2006: 348)이라고도 불리는 대안적인 창발론이 제안되었다. 특히 현재 대표적인 인과적 창발론자들 가운데 한 사람인 오코너는 과거에 수반 창발론을 옹호하였지만(O'Connor 1994), 최근에는 입장을 바꾸어 인과적 창발론을 주장하고 있다. 그는 기본적으로 한 시점에서의 창발적 속성이 이전의 상태(기저 속성과 창발적 속성을 모두 포함하는)에 의해 인과적으로 산출되는 통시적 관계에 집중하고 있다. (이는 그가 창발 속성과 기저 속성 사이의 공시적인 하향식 인과가 순환적일 수밖에 없다고 생각하기 때문은 아니다.)⁵³⁾ 그와 같은 입장에서 볼 때, 한 시점에서 창발 속성과 기저 속성 사이이 수반 관계가 성립하는지 묻는 것

53) 오코너가 취하는 전략은 다음과 같다. 그는 한 시점에서의 물리적 속성을 두 묶음으로 나눈다. 둘 중에서 한 묶음(P)는 창발적 속성(E)을 인과적으로 산출하고, 다른 한 묶음(P*)은 창발적 속성에 의해 인과적으로 영향을 받는다. 그러면 공시적 인과에 의한 순환성이 발생하지 않게 된다는 것이다(O'Connor and Wong 2005: 669). 공저자인 웡은 이후에 쓴 논문에서 같은 전략을 반복한 바 있다(Wong 2010: 9). 그런데 이러한 전략을 취하더라도 인과 관계가 전이적이라면 순환성이 발생할 것이다. 그리고 수반을 받아들이는 고전적인 창발론의 경우에 세 속성들(P, P*, E) 사이의 수반 관계가 어떻게 되는지도 의문이다. (수반 관계는 재귀성과 대칭성을 허용하기 때문에 큰 문제는 없을지도 모른다.) 내가 보기에 이들은 이러한 전략에 깊이 개입하고 있지 않다. 이들은 정말로 문제가 된다면 공시적 인과 관계를 기꺼이 버리려고 할 것이다. 그래도 어쨌든 공시적 인과 관계의 순환성이라는 문제를 해결하고자 한다면, 차라리 멈포드처럼 인과 관계에 대한 대안적인 모델을 채택하는 편이 나을 것이다(Mumford and Anjum 2011: Ch.5, Mumford 2013). 역량의 존재론을 주장하는 그는 인과 관계를 시간적으로 떨어져 있는 두 사건 사이의 관계로 보는 표준적인 모델을 거부하고 역량들이 서로 만나는 순간 시작되는 연속적 과정으로 본다. 이 과정에서 역량들은 상호적으로 표현되며, 대상들은 그것들이 가진 역량들을 얻거나 잃을 수 있다. 고전적인 사례에 이러한 모델을 적용하자면, 굴러가는 한 당구공은 정지해 있는 당구공과 접촉하는 그 순간 운동량이라는 서로의 역량을 주고 받는 인과적 과정에 돌입하는데, 그 과정은 두 물체가 완벽한 강체가 아니기 때문에 미세하게 찌그러지는 동안 지속되다가, 두 물체가 서로 떨어지는 순간 끝난다. 이러한 모델에서 원인과 결과 사이의 구별은 그것들이 인과적 과정의 시작과 끝을 의미하는 것이라면 시간적 선후 관계에 의해 설명될 수도 있지만, 그렇지 않은 경우에는 역량의 전달이 어디로부터 어디로 행해지는가에 의해 설명된다. 중요한 점은 이 대안적인 모델이 진정으로 공시적인 대칭적 인과 역시 허용한다는 점이다(Mumford and Anjum 2011: 117-19). 피드백 관계가 그러한 공시적인 대칭적 인과 관계의 대표적인 사례로 거론될 수 있다. 이러한 대안적인 모델을 채택한다면, 창발적 속성과 기저 속성 사이의 공시적 인과 관계가 순환적이라는 것은 문제가 되지 않는다. 다만 인과 관계를 이처럼 연속적 과정으로 보는 대안적인 모델을 채택할 경우, 인과적 창발론에서 주로 논하는 통시적 인과 관계 역시 다른 식으로 이해되어야 한다는 과제는 남는다.

은 결국 원인이 결과들을 필연화하는지를 묻는 것이다.

우리는 앞선 논의들을 통해 이미 법칙과 인과에 관한 특정한 견해들을 가지게 되었는데, 그것에 비추어 보면 수반 관계가 성립하지 않을 수도 있다는 점을 깨닫게 된다. 우리는 이미 단칭 인과가 그것을 포괄하는 법칙을 필요로 하지 않는다는 단칭주의를 받아들였고,⁵⁴⁾ 그것과 잘 어울리게 법칙은 역량에 대한 서술이라는 견해 역시 받아들였다. 이러한 견해에 따르면, 인과는 곧 특정한 상황 속에서 하나의 역량의 표현되는 것인데, 상황은 다시 다양한 역량들로 이루어져 있으므로, 결국 인과란 역량들의 상호적 표현이라고 간단히 말해질 수 있다. 역량이 제대로 표현되는지의 여부는 관련된 다른 역량들의 배열에 의존적이므로, 원인은 결과를 필연화하지 않는다. 앞서 보았듯이, 이런 그림에서 원인은 결과를 산출하는 경향성에 불과하다. 또, 관련된 모든 역량들의 배열이 고정되어 있다고 하더라도 비결정론이 참이라면 총체적 수반 역시 성립하지 않을 수 있다. 이전 시점에서의 세계가 현재 시점에서 같은 물리적 속성을 가지면서도 서로 다른 창발적 속성을 가지는 세계를 인과적으로 산출할 가능성이 있다면, 총체적 수반은 성립하지 않는다(O'Connor and Wong 2005: 667-68). 그리고 오늘날 우리는 비결정론을 진지하게 받아들여야 할 좋은 이유들을 가지고 있다.⁵⁵⁾ 따라서 인과 관계에 대해 이 글이 가지고 있는 그림에 비추어 볼 때, 인과적 창발론을 채택할 경우 국지적인 강수반은 물론이고 총체적 수반 역시 성립하지 않을 것이다.

인과적 창발론에서 위와 같이 수반이 실패할 가능성이 생겨나는 이유는 공시적인 공변 관계를 지정하는 엄격하고 결정론적인 창발 법칙을 받아들이지 않기 때문이다. 반대로 고전적인 창발론에서 창발 속성이 기저 속성에 법칙적 필연성에 의해 수반하는 이유는 창발 법칙이 엄격한 결정론적

54) 인과적 창발론자 가운데 한 사람으로 분류될 수 있는 험프리(Humphreys 1997: 13)도 단칭주의를 받아들이고 있다. 그는 역량에 대해 말하지는 않지만, 이 글과 마찬가지로 단칭 인과에서 원인이 결과를 필연화하지 못한다는 이유로 수반을 거부한다.

55) 비결정론이 완전히 배제된다면, 인과적 창발론에서도 총체적 수반은 성립한다. 하지만 이때의 결정론적인 세계는 물리주의자들이 생각하는 미시 물리적 존재자들에 의한 상향식 인과만이 존재하는 결정론적 세계가 아니라 온갖 하향식 인과들도 존재하는 특이한 결정론적 세계이다.

법칙이기 때문이다. 고전적인 창발론에서도 만약 두 속성 사이를 연결하는 창발 법칙이 비결정론이 법칙이거나 혹은 더 나아가 경향성에 불과하다면, 법칙적 필연성에 의한 수반 관계는 성립하지 않을 수 있다. 고전적인 창발론 자체는 이 가능성을 배제하지 못한다. 따라서 수반 관계는 창발론의 본질적인 요소가 아니다. 이는 놀라운 소식이 아니다. 웅이 지적하였듯이, 수반이 단지 두 속성 사이의 필연적 공변 관계만을 의미하는 것이라면, 심지어 데카르트식 실체 이원론에서도 두 실체 사이를 연결하는 모종의 결정론적 법칙이 있다는 전제 하에, 법칙적 필연성에 의한 수반 관계는 사소하게 성립할 수 있다(Wong 2010: 16).

고전적인 창발론과 인과적 창발론의 진정한 차이는 설명의 순서에서 공시적 관계가 먼저 와야 하는가, 아니면 통시적 관계가 먼저 와야 하는가에 있는 것으로 보인다. 그 효과는 만약 공시적 관점에서 수반이 성립할 경우 전자는 고유한 창발 법칙에 의해서만 수반 관계를 설명해야 하는 반면, 후자는 결정론적 인과 관계에 의해 설명할 수 있게 된다는 것이다. 창발 법칙을 도외시한다면, 공시적 설명과 통시적 설명 중에 어떤 것이 우선해야 하는가는 별도의 논의를 필요로 한다. 어쨌든 다시 한 번 강조되어야 할 점은 창발론에서 수반은 필수적인 요소가 아니며, 법칙적 충분성에 의해 창발 개념을 이해하는 것은 이미 특정한 입장을 취한 결과라는 것이다.

끝으로 창발론은 무언가 새로운 것이 생겨난다는 그 이름이 주는 선입견 때문에 종종 잘못 제기된 물음에 답변하도록 부당한 요구를 받기도 했다. 그 물음이란 흔히 ‘어떻게 단순한 물리적 입자들의 운동이 정신적 현상을 비롯한 존재론적으로 새로운 현상을 만들어내는가?’라는 식으로 제기된다. 하지만 다양한 속성들과 그것에 의해 개별화되는 다양한 실체들은 우리가 보다 기본적인 존재자들을 가지고 정말로 만들어내야 하는 것들이 아니다. 그것들은 어떤 설명 방식에서는 보다 기본적인 존재자들보다 나중에 올 수도 있겠지만, 실제로는 처음부터 이 현실에 이미 존재하고 있는 것들이다(Cartwright 1999: 33, O’Connor and Wong 2005: 660, Jaworski 2013: 160).

4.5. 결론

근본적 속성의 다수성을 인정하는 것은 일반화된 유형 부수현상론에 대한 가장 강력하고도 직접적인 대응이다. 그리고 실체 다원론은 근본적 속성의 다수성과 가장 어울리는 조합이다. 실체에 관한 이론들 가운데서도 실체를 유기적 통일성의 방식 자체로 보는 아리스토텔레주의적 실체 이론은 데카르트식의 순수 실체 이론이나 다른 실체 이론이 겪는 문제들을 겪지 않으면서 정신을 비롯한 다양한 존재자들이 어떤 식으로 존재하는지를 가장 그럴듯하게 보여줄 수 있는 이론이라는 점에서 평가절상되어야 할 것이다.

참 고 문 헌

- 데이빗슨, D., 1967/2012a, 「인과 관계」, 『행위와 사건』(배식한 옮김), 한길사.
- , 1969/2012b, 「사건의 개별화」, 『행위와 사건』(배식한 옮김), 한길사.
- , 1970/2012c, 「정신적 사건」, 『행위와 사건』(배식한 옮김), 한길사.
- , 1974/2012d, 「철학의 일종인 심리학」, 『행위와 사건』(배식한 옮김), 한길사.
- 루, M., 2010, 『형이상학 강의』(박제철 옮김), 아카넷.
- 백도형, 2014, 『심신 문제』, 아카넷.
- 험펠, C. G., 2014, 『자연 과학 철학』(곽강제 옮김), 서광사.
- Anscombe, G. E. M., 1971/1993, “Causality and Determination”, in E. Sosa and M. Tooley (eds.), *Causation*, Oxford University Press, pp. 88-104.
- Audi, R., 1993, “Mental Causation: Sustaining and Dynamic”, in Heil and Mele 1993, pp. 53-74.
- Baker, L. R., 1993, “Metaphysics and Mental Causation”, in Heil and Mele 1993, pp. 75-95.
- Bennett, K., 2003, “Why the Exclusion Problem Seems Intractable, and How, Just Maybe, to Tract It”, *Noûs* 37:3, pp. 471-497.
- , 2007, “Mental causation”, *Philosophy Compass* 2: 316-337.
- , 2008, “Exclusion again” In J. Kallestrup and J. Hohwy (eds.), *Being Reduced: New Essays on Reduction, Explanation, Causation*, Oxford University Press. pp. 280-305.
- Burge, T., 1993, “Mind-Body Causation and Explanation”, in Heil

- and Mele 1993, pp. 97-120.
- Carroll, J. W., 2009, "Anti-Reductionism", In H. Beebe, C. Hitchcock, and P. Menzies (eds.), *The Oxford Handbook of Causation*, Oxford University Press, pp. 279-298.
- Cartwright, N., 1983. *How the Laws of Physics Lie*, Oxford University Press.
- , 1989, *Nature's Capacities and Their Measurement*, Oxford University Press.
- , 1999, *The Dappled World: A Study of the Boundaries of Science*, Cambridge University Press.
- . 2008, "Reply to Ronal N. Giere", in Hartmann, Hoefer and Bovens 2008, Routledge. pp. 134-136.
- , 2010 "Natural laws and the closure of physics", In R. Y. Chiao, M. L. Cohen, A. J. Leggett, W. D. Phillips and C. L. Harper, Jr (eds.), *Visions of Discovery: New Light on Physics, Cosmology, and Consciousness*, Cambridge University Press, pp. 612-623.
- Cartwright, N. and J. Pemberton, 2013, "Aristotelian Powers: Without Them, What Would Modern Science Do?" in Greco and Gross 2013, pp. 93-112.
- Child, W., 1994, *Causality, Interpretation and the Mind*, Oxford University Press.
- Davidson, D., 1993, "Thinking Causes", in Heil and Mele 1993, pp. 3-17.
- , 2001, "Reply to Quine on Events" reprinted in *Essays on Actions and Events*, Clarendon Press. pp. 305-311.
- Dretske, F., 1989, "Reasons and Causes", *Philosophical Perspectives* 3, pp. 1-15.
- . 1993. "Mental Events as Structuring Causes of Behavior"

- in Heil and Mele 1993, pp. 121-136.
- Dupré, J., 2010, "How to be Naturalistic Without Being Simplistic in the Study of Human Nature", M. De Caro and D. Macarthur (eds.), *Naturalism and Normativity*, Columbia University Press, pp. 289-303.
- Gibb, S. C., 2013, "Introduction", Gibb, Lowe and Ingthorsson 2013, Oxford University Press. pp. 1-17.
- Gibb, S. C., E. J. Lowe, and R. D. Ingthorsson (eds.), 2013, *Mental Causation and Ontology*, Oxford University Press.
- Giere, R., 2008, "Models, Metaphysics and Methodology", in Hartmann, Hoefer and Bovens 2008, Routledge. pp. 123-133.
- Greco, J. and R. Gross (eds.), 2013, *Powers and Capacities in Philosophy: The New Aristotelianism*, Routledge.
- Hartmann S., C. Hoefer and L. Bovens, 2008, *Nancy Cartwright's Philosophy of Science*, Routledge.
- Hoefer, C., 2008, "For Fundamentalism", in Hartmann, Hoefer and Bovens 2008, Routledge. pp. 307-321.
- Heil, J. and A. Mele (eds.), 1993, *Mental Causation*, Oxford: Clarendon Press.
- Humphreys, P., 1997, "How Properties Emerge", *Philosophy of Science* 64:1, pp. 1-17.
- Jaworski, W., 2013, "Powers, Structures, and Minds", in Greco and Gross 2013, pp. 145-171.
- Kim, J., 1993a, "Can Supervenience and 'Non-Strict Laws' Save Anomalous Monism?", in Heil and Mele 1993, pp. 19-26.
- , 1993b, "The Non-Reductivist's Troubles with Mental Causation", in Heil and Mele 1993, pp. 189-210.
- Lowe, E. J., 2006, "Non-Cartesian Substance Dualism and the Problem of Mental Causation", *Erkenntnis* 65, pp. 5-23.

- , 2012, “A neo-Aristotelian substance ontology: neither relational nor constituent” In T. E. Tahko (ed.), *Contemporary aristotelian metaphysics*, Cambridge University Press, pp. 229–248.
- McLaughlin, B. P., 1989, “Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical”, *Philosophical Perspectives* 3, pp. 109–35.
- , 1993, “On Davidson's Response to the Charge of Epiphenomenalism”, in Heil and Mele 1993, pp. 27–40.
- Moore, M. S., 2009, *Causation and Responsibility: An Essay in Law, Morals, and Metaphysics*, Oxford University Press.
- Mumford, S., 2013, “The Power of Power”, in Greco and Gross 2013, pp. 9–24.
- Mumford, S. and R. L. Anjum, 2011, *Getting Causes from Powers*, Oxford University Press.
- O'Connor, T., 1994, “Emergent Properties”, *American Philosophical Quarterly* 31, pp. 91–104.
- O'Connor T., and H. Y. Wong, 2005, “Metaphysics of Emergence”, *Noûs* 39:4, pp. 658–678
- Putnam, H., 1999, *The Threefold Cord: Mind, Body, and World*, Columbia University Press.
- Pruss, A. R., forthcoming, “Aristotelian Forms and Laws of Nature”, URL = <http://www.alexanderpruss.com/papers/Forms.html>.
- Robb, D. and J. Heil, 2014, “Mental Causation”, in *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2014/entries/mental-causation/>.

- Robinson, H., "Substance", *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), Edward N. Zalta (ed.), URL = [<http://plato.stanford.edu/archives/spr2014/entries/substance/>](http://plato.stanford.edu/archives/spr2014/entries/substance/)
- Schaffer, J.. 2003 "Is There a Fundamental Level?", *Noûs*, 37: 498-517.
- , 2009 "On What grounds What" in D. Manley, D. J. Chalmers and R. Wasserman (eds.), *Metametaphysics: New Essays on the Foundations of Ontology*, Oxford University Press, pp. 347-383.
- Schneider, S. 2010, "Why property dualists must reject substance physicalism", *Philosophical Studies*, DOI 10.1007/s11098-010-9618-9.
- Shoemaker, S., 2013, "Physical Realization without Preemption", in Gibb, Lowe, and Ingthorsson 2013, Oxford University Press, pp. 35-57.
- Stump, E., 2013, "Emergence, Causal powers, and Aristotelianism in Metaphysics", In Greco and Gross 2013, pp. 48-68.
- Van Gulick, R., 1993, "Who's in Charge Here? And Who's Doing All the Work?", in Heil and Mele 1993, pp. 232-256.
- Woodward, J., 2014, "Scientific Explanation", *The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition), Edward N. Zalta (ed.), URL = [<http://plato.stanford.edu/archives/win2014/entries/scientific-explanation/>](http://plato.stanford.edu/archives/win2014/entries/scientific-explanation/).
- Wong, H. Y., 2006, "Emergents from Fusion", *Philosophy of Science* 73:3, pp. 345-367.
- , "The Secret Lives of Emergents", in A. Corradini and T. O'Connor (eds.), *Emergence in Science and Philosophy*, Routledge, pp. 7-24.

Yang, E., 2015, “The compatibility of property dualism and substance materialism”, *Philosophical Studies*, DOI 10.1007/s11098-015-0465-6.

Abstract

Substance Pluralist Solution to Mental Causation Problem

SeungTae Jeon

Department of Philosophy

The Graduate School

Seoul National University

The most prevalent answer to mental causation problem can be summed up in a position which is non-reductive physicalism. It is the thesis that denies mental type or property can be reduced into physical type or property, yet holds fast to physicalism. Roughly speaking, there are two versions of non-reductive physicalism. The one accepts token-identity, while the other rejects it. It seems that the later was suggested to avoid the charge of type-epiphenomenalism of the mental because the former has been criticized for that. however, non-reductive physicalism which denies token-identity seems to be in the difficult situation where it has to say mental token has no distinctive causal power but it is somehow distinguished from physical token. As we naturally accept the idea that properties are individuated by their distinctive causal powers, non-reductive physicalism cannot give the satisfactory answer to mental

causation problem.

The failure of non-reductive physicalism is due to the self-evidently held principle named 'causal closure of the physical' (or simply 'causal closure'). As long as we accept causal closure, non-physical properties find difficulty in possessing independent causal powers even at token level. Therefore, if we are to wholly admit the causal powers of non-physical properties, causal closure should be gave up. Non-reductive physicalists present unsatisfactory solutions, because they believe that causal closure is apparent enough to be exempted from being abandoned.

But it is the argument of this thesis that causal closure is not self-evident, and if we are to give fully satisfactory answer to mental causation problem, it should be rejected. Especially, I think we can learn from pioneering works of Nancy Cartwright and others in philosophy of science that the current theory and practice of sciences can be explained better without presupposing causal closure rather than with it. If it is indeed so, there is no need to excessively adhere to the causal closure.

If we reject causal closure and accept the fundamentality of mental property, a large variety of options to solve mental causation problem will be given. Among them, I tried to show Aristotelian hylomorphic substance pluralism which can be understood as a kind of emergentism is more suitable in many ways than some other major alternatives.

keywords : mental causation, non-reductive physicalism, causal closure, pluralism, emergentism, Cartwright, Aristotle

Student Number : 2013-22786